

ON EFFECTIVE EQUIDISTRIBUTION FOR HIGHER STEP NILFLOWS

LIVIO FLAMINIO AND GIOVANNI FORNI

ABSTRACT. The main goal of this paper is to obtain optimal estimates on the speed of equidistribution of nilflows on higher step nilmanifolds. Under a Diophantine condition on the frequencies of the toral projection of the flow, we prove that for almost all points on the nilmanifold orbits become equidistributed at polynomial speed with exponent which decays quadratically as a function of the number of steps. The main novelty is the introduction of new techniques of renormalization (rescaling) in absence of a truly recurrent renormalization dynamics. Quantitative equidistribution estimates are derived from bounds on the scaling of invariant distributions (in Sobolev norms) and on the geometry of the nilmanifold under the rescaling.

CONTENTS

1. Introduction	2
2. Quasi-Abelian nilpotent flows and Weyl sums	5
2.1. Quasi-Abelian nilpotent Lie algebras and groups	5
2.2. Quasi-Abelian nilmanifolds and flows	8
2.3. Weyl sums as ergodic integrals	11
3. A Sobolev trace theorem	13
3.1. Width of nilpotent orbits	14
3.2. Sobolev a priori bounds	15
3.3. Nilpotent averages	17
4. The cohomological equation	18
4.1. Irreducible unitary representations	18
4.2. A priori estimates	22
4.3. The renormalisation group	25
5. Average width estimates	28
5.1. Almost periodic points	29
5.2. Expected width bounds	32
5.3. Diophantine estimates	37
5.4. Width estimates along orbit segments	43
6. Bounds on ergodic averages	44
6.1. Coboundary estimates for rescaled bases	46
6.2. Bounds on ergodic averages in a irreducible sub-representation	48
6.3. General bounds on ergodic averages	53
References	59

Date: July 15, 2014.

1991 Mathematics Subject Classification. 37A17, 37A45, 11K36, 11L15.

Key words and phrases. Nilflows, Cohomological Equations, Ergodic averages.

1. INTRODUCTION

In this paper we prove estimates on the speed of ergodicity for a class of nilflows on higher step nilmanifolds, under Diophantine conditions on the frequencies of their toral projections. By the classical theory, nilflows with minimal, hence uniquely ergodic, toral projection are uniquely ergodic. Their ergodic theory is closely related to questions in number theory, in particular the problem of bounds on exponential sums along polynomial sequences, known as *Weyl sums*. By a relatively recent far reaching generalisation by B. Green and T. Tao [GT12] of classical results and methods, all orbits of Diophantine nilflows on any nilmanifold become equidistributed at polynomial speed, but the exponent in their theorem is far from optimal and presumably decays exponentially as a function of the number of steps of the nilmanifold.

We are especially concerned with the optimal speed of equidistribution for nilmanifolds of higher step. Our main result proves equidistribution at a polynomial speed with exponent which decays *quadratically* as a function of the number of steps. However, we only establish our result for *almost all points* on the nilmanifold. In other terms, we prove a rather sharp result on quantitative ergodicity, but for reasons that will be explained below, we are unable to prove an effective *unique* ergodicity theorem. Our result can be better appreciated by comparing its application to Weyl sums, stated below, with recent results proved by T. D. Wooley [Woo12] with methods of analytic number theory. In fact, we derive a virtually identical bound on the growth of Weyl sums for polynomials of higher degree under a comparable (but somewhat stronger) Diophantine condition on the leading coefficient. However, our result only holds for almost all choices of coefficients of lower degree.

We do not consider general nilmanifolds, but only a class of them which we call *quasi-Abelian*. This class is in a sense the simplest class of nilmanifolds of arbitrarily high step. A quasi-Abelian nilpotent group is a nilpotent group which contains an Abelian normal subgroup of codimension one. This class of quasi-Abelian nilpotent groups is chosen since on the one hand their irreducible unitary representations, which can be described as an application of Kirillov theory, are particularly simple, and on the other hand this class contains groups of arbitrarily high step, which allow us to derive results on Weyl sums for polynomials of arbitrarily high degree. There is no reason in principle that prevents a generalisation to arbitrary nilflows on arbitrary nilmanifolds, except that require estimates in representations would be very complicated and difficult to carry out.

Let $G_n^{(k)}$ denote a quasi-Abelian k -step nilpotent group on $n + 1$ generators, let $\Gamma_n^{(k)} \subset G_n^{(k)}$ be a lattice and let $M_n^{(k)} := \Gamma_n^{(k)} \backslash G_n^{(k)}$ denote the corresponding nilmanifold. Since the Abelianisation $G_n^{(k)} / [G_n^{(k)}, G_n^{(k)}]$ of the group G is isomorphic to \mathbb{R}^{n+1} , there is a natural projection $M_n^{(k)} \rightarrow \bar{M}_n^{(k)}$ onto an $n + 1$ -dimensional torus. By the classical theory, a nilflow $M_n^{(k)}$ is uniquely ergodic if and only if the projected toral flow on $\bar{M}_n^{(k)}$ has rationally independent frequencies.

Effective equidistribution results require a Diophantine condition on the frequencies. We formulate below our condition (see Definition 5.8). Let $\|\cdot\|_{\mathbb{Z}}$ denote the distance from the nearest integer. For any $\alpha := (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$, for any $N \in \mathbb{N}$

and for every $\delta > 0$, let

$$\mathcal{R}_\alpha^{(n)}(N, \delta) = \{r \in [-N, N] \cap \mathbb{Z} \setminus \{0\} \mid \max_{1 \leq i \leq n} \|r\alpha_i\|_{\mathbb{Z}} \leq \delta^{\frac{1}{n}}\}.$$

For every $\nu \geq 1$, let $D_n(\nu) \subset (\mathbb{R} \setminus \mathbb{Q})^n$ be the subset defined as follows: $\alpha \in D_n(\nu)$ if and only if there exists a constant $C(\alpha) > 0$ such that, for all $N \in \mathbb{N}$ and for all $\delta > 0$,

$$\#\mathcal{R}_\alpha^{(n)}(N, \delta) \leq C(\alpha) \max\{N^{1-\frac{1}{\nu}}, N\delta\}.$$

For a single frequency the above Diophantine condition is a consequence of the well-known following Diophantine condition. A number $a \in \mathbb{R} \setminus \mathbb{Q}$ is called Diophantine of exponent $\nu \geq 1$ if there exists a constant $c(a) > 0$ such that the following bound holds:

$$\|Na\|_{\mathbb{Z}} \geq \frac{c(a)}{N^\nu}, \quad \text{for all } N \in \mathbb{N} \setminus \{0\}.$$

By an elementary argument based on continued fractions, it can be proved that our set $D_1(\nu)$ introduced above contains all Diophantine irrational numbers of Diophantine exponent $\nu \geq 1$, according to the above classical definition. In higher dimension $n \geq 2$ our set $D_n(\nu)$ contains the set of simultaneously Diophantine vectors of sufficiently small exponent, hence we can prove that the set $D_n(\nu)$ has full measure for sufficiently large $\nu \geq 1$ (see Lemma 5.12).

Our main result is the following bound on the speed of convergence of ergodic averages along almost all orbits of Diophantine quasi-Abelian nilflows.

Theorem 1.1. *Let (ϕ_α^t) be a nilflow on a k -step quasi-Abelian nilmanifold $M_n^{(k)}$ on $n+1$ generators such that the projected toral flow $(\bar{\phi}_\alpha^t)$ is a linear flow with frequency vector $\alpha := (1, \alpha_1, \dots, \alpha_n) \in \mathbb{R} \times \mathbb{R}^n$. Under the assumption that the vector $\alpha' := (\alpha_1, \dots, \alpha_n) \in D_n(\nu)$ for some $\nu \leq k/2$, there exists a (Sobolev) norm $\|\cdot\|$ on the space $C^\infty(M_n^{(k)})$ of smooth function on $M_n^{(k)}$ and for every $\epsilon > 0$ there exists a positive measurable function $K_\epsilon \in L^p(M_n^{(k)})$ for all $p \in [1, 2)$, such that the following bound holds. For every smooth zero-average function $f \in C^\infty(M_n^{(k)})$, for almost every $x \in M_n^{(k)}$ and for every $L \geq 1$,*

$$\left| \frac{1}{L} \int_0^L f \circ \phi_\alpha^t(x) dt \right| \leq K_\epsilon(x) L^{-\frac{2}{3(k+2n-2)(k-1)} + \epsilon} \|f\|.$$

The above theorem is best appreciated by its main corollary on Weyl sums, in comparison with available results proved by analytic number theory. We recall that given a polynomial $P_k(N)$ of degree $k \geq 2$, written as $P_k(N) := \sum_{j=0}^k a_j N^j$ the corresponding Weyl sums are the exponential sums

$$W(a_k, \dots, a_0, N) := \sum_{n=0}^{N-1} \exp(2\pi i P_k(n)).$$

By the well-known relation between Weyl sums and nilflows, we derive the following bound.

Corollary 1.2. *Let $a_k \in \mathbb{R} \setminus \mathbb{Q}$ be a Diophantine number of exponent $\nu \leq k/2$. For every $\epsilon > 0$, there exists a measurable positive function $K_\epsilon \in L^p(\mathbb{T}^{k-2})$, for all $p \in [1, 2)$, such that the following bound holds. For all $a_0, a_1 \in \mathbb{R}^2$, for almost all $(a_2, \dots, a_{k-1}) \in \mathbb{R}^{k-2}$ and for every $L \geq 1$,*

$$|W(a_k, \dots, a_0, N)| \leq K_\epsilon(a_2, \dots, a_{k-1}) N^{1 - \frac{2}{3k(k-1)} + \epsilon}.$$

As we have anticipated above, this result should be compared with bounds proved by T. D. Wooley (see in particular Theorem 1.5 in [Woo12]). From Wooley’s theorem, one can easily derive a *uniform* bound on Weyl sums with essentially the same exponent as above, but for *all* $(a_0, \dots, a_{k-1}) \in \mathbb{R}^k$, also under a Diophantine condition of exponent $\nu \leq k - 1$. Wooley’s theorem comes as a refinement and sharpening of techniques that have been developed over the span of a century, and especially since Vinogradov’s contribution in the 30’s, and in Wooley’s words come “within a stone’s throw of the sharpest possible bounds” (for large degree $k \geq 8$, otherwise the classical Weyl’s bound is still unsurpassed in general).

Our approach is significantly different from the methods of analytic number theory and the circle methods of Wooley’s [Woo12] as well as from the classical methods based on induction on the number of steps and on Van der Corput lemma greatly refined recently in the work of Green and Tao [GT12]. In their work quantitative equidistribution with polynomial speed is proved for general nilpotent sequences. However, there is no effort to determine the optimal exponent as a function of the number of steps. In fact, since their methods are a generalisation of the Weyl’s method, it is reasonable to expect that the best exponent available in their work would decay *exponentially* with the number of steps (we recall that the classical Weyl bound on Weyl sums holds with exponent $1 - 1/2^{k-1}$ for polynomial sequences of degree $k \geq 2$).

In this paper we generalise the renormalisation method of our earlier work [FF06] on Heisenberg nilflows (the 2-step nilpotent case, which corresponds to polynomial sequences of degree 2). Our main goal is to develop an approach which is not restricted to nilflows and can be applied to quantitative equidistribution problems for more general parabolic flows. The method is based on estimates on the scaling of invariant distributions for the flow under a deformation of the nilmanifold. In the Heisenberg case, it is possible to define a deformation given by a one-parameter group of automorphism of the Lie algebra, which implies that the deformation group induces a renormalisation group action on a suitable moduli space. This is not surprising since it has been known for a long time that quadratic polynomial sequences (as well as linear ones) have self-similarity properties.

In the higher step case, we were unable to define an effective renormalisation group dynamics and our deformation does not come from a group of Lie algebra automorphisms. As a consequence, it does not induce a recurrent flow on a moduli space. However, quantitative equidistribution estimates can still be derived from bounds on the scaling of invariant distributions (in Sobolev norms) and on the geometry of the nilmanifold under the deformation. Given that no (recurrent) renormalisation is available, the task of proving geometric bounds is in fact the most delicate part of the argument. Our proof is based on average estimates and on a Borel-Cantelli argument, which explains why our geometric bounds, and consequently our equidistribution results, only holds almost everywhere. In fact, the deformation is chosen in a way that optimises the scaling of invariant distributions. Sobolev estimates on the scaling of invariant distributions can be proved by an analysis of the cohomological equation and of invariant distributions in every irreducible representation of the (quasi-Abelian) Lie group (the quasi-Abelian case is in fact much simpler than the general case treated in [FF07] and allows us to prove explicit sharp bounds). This analysis leads to the polynomial decay of ergodic averages (with the exponent given in our main theorem) for all “good” points for

which uniform bounds on the degeneration of the geometry hold. It is a plausible conjecture that in fact under a Diophantine condition on the nilflow *all* points of the nilmanifold, not just almost all, are “good”.

The degeneration of the geometry at a given point on the nilmanifold is measured in terms of a notion of *average width* of an orbit segment of a nilflow, with respect to a basis of the Lie algebra (see Definition 3.5). This notion arises from a new version of the Sobolev trace theorem adapted to Sobolev estimates on orbit segments of flows (see Theorem 3.10). From a dynamical standpoint, the average width is a measure of the frequency of close returns along an orbit segment. Roughly, the width of an orbit segment is the maximal transverse volume of a rectangular tubular neighbourhood, measured with respect to a given, possibly deformed, transverse metric. The inverse of the square root of the width bounds the constant in the Sobolev trace theorem, which provides an a priori Sobolev bound for the distribution given by an orbit segment. The average width is an averaged version of the width, with the average taken along the orbit segment itself. The tubular neighbourhood is allowed to have a rectangular cross section of variable transverse volume. The average width is defined as the reciprocal of the average along the orbit segment of the reciprocal of the transverse area of the tubular neighbourhood. The point is that if the very close returns of the orbit segment are not too frequent then the average width can still be (uniformly) bounded, while the width may be arbitrarily large. We prove that the average width still gives an upper bound for the constant in the Sobolev trace theorem.

The paper is organised as follows. In section 2 we define quasi-Abelian groups, nilmanifolds and nilflows and recall the well-known relations between Weyl sums and ergodic averages of nilflows. In section 3 we introduce the notion of average width and prove a Sobolev trace theorem. In section 4 we carry out Sobolev estimates on solutions of the cohomological equation and on invariant distributions as an application of Kirillov theory of unitary representations of nilpotent groups. In section 5 we prove bounds on the average width of orbit segment of nilflows, in mean over the initial point of the orbit. Finally, in section 6 we prove an effective equidistribution theorem for “good” points and derive that “good” points form a set of full measure, by a Borel-Cantelli argument based on the estimates in mean on the average width. Estimates on Weyl sums (for almost all lower degree coefficients) then follow from our equidistribution theorem.

Acknowledgements Livio Flaminio was supported in part by the Labex CEMPI (ANR-11-LABX-07). Giovanni Forni was supported by NSF grant DMS 1201534. This work was completed at the Isaac Newton Institute in Cambridge, UK. The authors wish to thank the Institute and the organisers of the programme *Interactions between Dynamics of Group Actions and Number Theory* for their hospitality.

2. QUASI-ABELIAN NILPOTENT FLOWS AND WEYL SUMS

In this section we introduce quasi-Abelian nilpotent Lie algebras and groups and collect basic material on their structure. We recall classical Weyl sums and a well-known reduction of Weyl sums to ergodic integrals of quasi-Abelian nilpotent flows [Fur81].

2.1. Quasi-Abelian nilpotent Lie algebras and groups.

2.1.1. Quasi-Abelian nilpotent Lie algebras. In this paper a nilpotent Lie algebra \mathfrak{g} will be called *quasi-Abelian* if it has a *maximal* Abelian ideal \mathfrak{a} of codimension one. We denote by a the dimension of \mathfrak{a} . Any quasi-Abelian nilpotent Lie algebra admits bases

$$(1) \quad \{\xi, \eta_1^{(1)}, \dots, \eta_{i_1}^{(1)}, \dots, \eta_1^{(n)}, \dots, \eta_{i_n}^{(n)}\},$$

such that the only non-trivial commutation relations are of the form

$$(2) \quad [\xi, \eta_i^{(m)}] = \eta_{i+1}^{(m)}, \quad \text{for all } m = 1, \dots, n, \text{ and } i = 1, \dots, i_m - 1.$$

Such bases can be constructed as follows. Let $\xi \notin \mathfrak{a}$, let

$$\mathfrak{a} = \mathfrak{a}^{(1)} \oplus \dots \oplus \mathfrak{a}^{(n)},$$

be the splitting of the vector space \mathfrak{a} into generalised eigenspaces of the linear map $\text{ad}(\xi) : \mathfrak{a} \rightarrow \mathfrak{a}$ and for every $m = 1, \dots, n$ let

$$(3) \quad \{\eta_1^{(m)}, \dots, \eta_{i_m}^{(m)}\} \subset \mathfrak{a}^{(m)}$$

be a Jordan basis for the linear map $\text{ad}(X) : \mathfrak{a}^{(m)} \rightarrow \mathfrak{a}^{(m)}$. For this reason, bases satisfying the commutation relations (2) will be called *Jordan bases*.

Any quasi-Abelian nilpotent Lie algebra with a Jordan basis of the form (1), (2) is generated (as a Lie algebra) by the system $\{\xi, \eta_1^{(1)}, \dots, \eta_1^{(n)}\}$, hence it has $n + 1$ generators, and it has an Abelian ideal \mathfrak{a} of dimension $a = i_1 + \dots + i_n$ generated (as a vector space) by the system in formula (3), hence the Lie algebra \mathfrak{g} has dimension $a + 1$. Finally, the Lie algebra is k -step nilpotent for

$$k := \max\{i_1, \dots, i_n\}.$$

In this paper we are interested in quasi-Abelian nilpotent Lie algebras of step $k \geq 3$.

Notation 2.1. It will be convenient to consider the set of indices

$$J := \{(m, i) \mid m = 1, \dots, n, i = 1, \dots, i_m\}.$$

endowed with the lexicographic order. By J^- we denote the ordered subset $\{(m, j) \in J \mid i \leq i_m - 1\}$.

Definition 2.2. An ordered basis $(X, Y) := (X, \dots, Y_i^{(m)}, \dots)_{(m,i) \in J}$ of the quasi-Abelian Lie algebra \mathfrak{g} is a *generalised Jordan basis* if $X \notin \mathfrak{a}$, $Y_i^{(m)} \in \mathfrak{a}$ for all $(m, i) \in J$ and, for some strictly positive reals $c = (c_i^{(m)})_{(m,i) \in J^-}$, the following commutation relations hold true

$$[X, Y_i^{(m)}] = c_{i+1}^{(m)} Y_{i+1}^{(m)}, \quad \text{for all } (m, i) \in J^-,$$

(all other commutators being equal to zero). The constants c are called the *structural constants of the basis*.

2.1.2. Quasi-Abelian nilpotent Lie groups. A nilpotent group G will be called a *quasi-Abelian* k -step nilpotent Lie group (on $n + 1$ generators) if it is a simply connected, connected Lie group whose Lie algebra \mathfrak{g} is a quasi-Abelian k -step nilpotent Lie algebra (on $n + 1$ generators), as above. A quasi-Abelian nilpotent Lie group G has an Abelian normal subgroup A of codimension one, namely the exponential of the codimension one, Abelian ideal \mathfrak{a} of \mathfrak{g} .

The k -step quasi-Abelian nilpotent groups G on $n + 1$ generators have also another description. Let $(i_1, \dots, i_n) \in \mathbb{Z}_+^n$ be positive integers such that $k = \max\{i_1, \dots, i_n\}$ and let $a = i_1 + \dots + i_n$.

For any $j \in \mathbb{Z}_+$, let $h_j : \mathbb{R} \rightarrow \text{Aut}(\mathbb{R}^j)$ be the (unique) one-parameter group of automorphisms of \mathbb{R}^j such that

$$(4) \quad h_j(1)(s_1, \dots, s_j) = (s_1, s_2 + s_1, \dots, s_i + s_{i-1}, \dots, s_j + s_j)$$

and let $h : \mathbb{R} \rightarrow \text{Aut}(\mathbb{R}^a)$ be the product one-parameter group

$$(5) \quad h = h_{i_1} \times \dots \times h_{i_n} \quad \text{on } \mathbb{R}^a = \mathbb{R}^{i_1} \times \dots \times \mathbb{R}^{i_n}.$$

Let G be the twisted product $\mathbb{R} \ltimes_h \mathbb{R}^a$. We can view G as an algebraic subgroup of the real algebraic group $\text{GL}_d(\mathbb{R}) \ltimes \mathbb{R}^a$. Since $h(\mathbb{Z}) \subset \text{Aut}(\mathbb{Z}^a)$, the twisted product $\Gamma := \mathbb{Z} \ltimes_{h|_{\mathbb{Z}}} \mathbb{Z}^a$ is a well-defined, Zariski dense, discrete subgroup of G , hence a lattice of G . It is generated by elements

$$x, y_1^{(1)}, \dots, y_{i_1}^{(1)}, \dots, y_1^{(n)}, \dots, y_{i_n}^{(n)},$$

such that the only non-trivial commutation relations are

$$xy_i^{(m)}x^{-1} = \begin{cases} y_i^{(m)}y_{i+1}^{(m)}, & \text{for } 1 \leq i < i_m \\ y_{i_m}^{(m)}, & \text{for } i = i_m. \end{cases}$$

(We have taken for x the element $(1, (0, \dots, 0)) \in \Gamma$ which acts by conjugation on \mathbb{Z}^a by the automorphism $h(1)$ defined in (4) and (5), and for elements $y_1^{(1)}, \dots, y_{i_1}^{(1)}, \dots, y_1^{(n)}, \dots, y_{i_n}^{(n)}$ the elements of the standard basis $(0, (1, 0, \dots, 0)), \dots, (0, (0, \dots, 0, 1))$ of $\{0\} \times \mathbb{Z}^a$). The codimension one, Abelian normal subgroup $A \subset G$ is generated by the elements

$$y_1^{(1)}, \dots, y_{i_1}^{(1)}, \dots, y_1^{(n)}, \dots, y_{i_n}^{(n)}.$$

Let \mathfrak{g} be the Lie algebra of G and let $\log : G \rightarrow \mathfrak{g}$ the inverse of the exponential map $\exp : \mathfrak{g} \rightarrow G$. The elements

$$(6) \quad \xi := \log x \quad \tilde{\eta}_i^{(m)} := \log y_i^{(m)}, \quad (m, i) \in J,$$

form a basis of \mathfrak{g} and satisfy the commutation relations

$$(7) \quad [\xi, \tilde{\eta}_j^{(m)}] = \sum_{i=j+1}^{i_m} \frac{(-1)^{i-j-1}}{i-j} \tilde{\eta}_i^{(m)}, \quad (m, j) \in J^-,$$

all other commutators being equal to zero. We obtain a Jordan basis defining by induction

$$(8) \quad \eta_1^{(m)} = \tilde{\eta}_1^{(m)}, \quad \eta_{i+1}^{(m)} = [\xi, \eta_i^{(m)}], \quad (m, i) \in J^-.$$

Thus \mathfrak{g} is a quasi-Abelian k -step nilpotent Lie algebra on $n + 1$ generators, hence G is a quasi-Abelian k -step nilpotent Lie group on $n + 1$ generators.

Clearly, for all $m = 1, \dots, n$ there exists strictly upper triangular rational matrices $R^{(m)}, S^{(m)} \in M_{i_m}(\mathbb{Q})$ such that

$$(9) \quad \eta_j^{(m)} = \tilde{\eta}_j^{(m)} + \sum_{i=j+1}^{i_m} R_{ij}^{(m)} \tilde{\eta}_i^{(m)}$$

$$(10) \quad \tilde{\eta}_j^{(m)} = \eta_j^{(m)} + \sum_{i=j+1}^{i_m} S_{ij}^{(m)} \eta_i^{(m)}.$$

for all $j = 1, \dots, i_m - 1$. Thus via the formulas (10) and by taking exponentials, we can associate a lattice of the quasi-Abelian nilpotent Lie group G to each Jordan basis $(\xi, \eta) = (\xi, \dots, \eta_i^{(m)}, \dots)$ of its Lie algebra \mathfrak{g} .

Henceforth we shall assume that we have fixed once and for all a Jordan basis $(\xi, \eta) = (\xi, \dots, \eta_i^{(m)}, \dots)$ of the Lie algebra \mathfrak{g} and we shall define Γ to be the lattice generated by the system

$$(11) \quad \{x := \exp \xi, \dots, y_i^{(m)} := \exp \tilde{\eta}_i^{(m)}, \dots\},$$

where the elements $\tilde{\eta}_i^{(m)} \in \mathfrak{g}$ are given by the formulas (10) and satisfy the commutation relations in formula (7).

2.2. Quasi-Abelian nilmanifolds and flows.

2.2.1. *Quasi-Abelian nilmanifolds.* Since by construction the subgroup Γ is discrete and Zariski dense in G the quotient $\Gamma \backslash G$ is a compact nilmanifold.

Definition 2.3. The quotient $M = \Gamma \backslash G$ will be called a *quasi-Abelian k -step nilmanifold on $n + 1$ generators*.

Observe that for any Jordan basis (ξ, η) the centre $\mathfrak{z}(\mathfrak{g})$ of a quasi-Abelian k -step nilpotent Lie algebra \mathfrak{g} is spanned by the system $\{\eta_{i_1}^{(1)}, \dots, \eta_{i_n}^{(n)}\}$ and therefore the system $(\xi, \dots, \eta_i^{(m)}, \dots)$, with $(m, i) \in J^-$, projects onto a Jordan basis of the Lie algebra $\mathfrak{g}' := \mathfrak{g}/\mathfrak{z}(\mathfrak{g})$, which is quasi-Abelian $(k - 1)$ -step nilpotent on $n' \leq n$ generators. At the group level, the centre $Z(G)$ of G , that is the group

$$\{\exp(t_1 \eta_{i_1}^{(1)} + \dots + t_n \eta_{i_n}^{(n)})\}_{(t_1, \dots, t_n) \in \mathbb{R}^n}$$

meets the lattice Γ into the subgroup generated by the system

$$\{\exp(\eta_{i_1}^{(1)}), \dots, \exp(\eta_{i_n}^{(n)})\},$$

which is the centre $Z(\Gamma)$ of Γ . Hence $G' := G/Z(G)$ is a quasi-Abelian $(k - 1)$ -nilpotent Lie group on $n' \leq n$ generators and the subgroup $\Gamma' := \Gamma/Z(\Gamma)$ is a lattice in G . In fact, the elements

$$x, y_1^{(1)}, \dots, y_{i_1-1}^{(1)}, \dots, y_1^{(n)}, \dots, y_{i_n-1}^{(n)}$$

project onto generators of the lattice Γ' in G' . The above discussion implies that a quasi-Abelian k -step nilmanifold M has a structure of toral bundle over a quasi-Abelian $(k - 1)$ -step nilmanifold M' the fibres of this fibration being the orbits of the right action of the centre $Z(G)$ on M .

We introduce two other important fibrations of the nilmanifold M . Since the Abelianisation $G/[G, G]$ of the group G is isomorphic to \mathbb{R}^{n+1} and contains the subgroup $\Gamma/[\Gamma, \Gamma]$ as a co-compact lattice we obtain the following fibration

$$(12) \quad 0 \rightarrow \mathbb{T}^{a-n} \rightarrow M \xrightarrow{\text{pr}_1} M_1 \approx \mathbb{T}^{n+1} \rightarrow 0.$$

Another fibration arises from the canonical homomorphism $G \rightarrow G/A \approx \langle \exp \xi \rangle$; passing to the quotient by the corresponding lattices, we see that M is a torus bundle over a circle with monodromy given by the map $h(1)$ of formulas (4) and (5), that is,

$$(13) \quad 0 \rightarrow \mathbb{T}^a \rightarrow M \xrightarrow{\text{pr}_2} M_2 \approx \mathbb{T}^1 \rightarrow 0.$$

For all $m = 1, \dots, n$ we denote by $\mathbb{T}_0^{i_m} \subset M$ the i_m -dimensional torus

$$\Gamma \exp(s_1^{(m)} \tilde{\eta}_1^{(m)} + \dots + s_{i_m}^{(m)} \tilde{\eta}_{i_m}^{(m)}), \quad (s_1^{(m)}, \dots, s_{i_m}^{(m)}) \in \mathbb{R}^{i_m}.$$

By construction the fibre \mathbb{T}_0^a of the fibration (13) above the coset of the identity has a product structure

$$(14) \quad \mathbb{T}_0^a \approx \mathbb{T}_0^{i_1} \times \dots \times \mathbb{T}_0^{i_n}.$$

For $m = 1, \dots, n$, let us denote $\mathbf{s}^{(m)} := (s_1^{(m)}, \dots, s_{i_m}^{(m)}) \in (\mathbb{R}/\mathbb{T})^{i_m}$ and $\mathbf{s} = (\mathbf{s}^{(1)}, \dots, \mathbf{s}^{(n)}) \in (\mathbb{R}/\mathbb{Z})^a$. The map

$$\mathbf{s}^{(m)} \in (\mathbb{R}/\mathbb{Z})^{i_m} \mapsto \Gamma \exp\left(\sum_{i=0}^{i_m} s_i^{(m)} \tilde{\eta}_i^{(m)}\right) \in \mathbb{T}_0^{i_m}$$

is a diffeomorphism, and so is the map

$$\mathbf{s} \in (\mathbb{R}/\mathbb{Z})^a \mapsto \Gamma \exp\left(\sum_{(m,i) \in J} s_i^{(m)} \tilde{\eta}_i^{(m)}\right) \in \mathbb{T}_0^a.$$

Points on the tori $\mathbb{T}_0^{i_m}$ and $\mathbb{T}_0^a \subset M$ will be denoted by their coordinates $\mathbf{s}^{(m)} \in (\mathbb{R}/\mathbb{Z})^{i_m}$ and $\mathbf{s} \in (\mathbb{R}/\mathbb{Z})^a$.

By the definition (11) of the lattice Γ the left (and right) invariant volume form vol_{a+1} on G , normalised by the condition $\text{vol}_{a+1}(\xi, \dots, \tilde{\eta}_i^{(m)}, \dots) = 1$ pushes down to a right-invariant volume form on ω on M , whose density yields a right-invariant *probability* measure \mathcal{L} on M . Since the formulas (9) and (10) imply that the wedge products $\xi \wedge \dots \wedge \tilde{\eta}_i^{(m)} \wedge \dots$ and $\xi \wedge \dots \wedge \eta_i^{(m)} \wedge \dots$ coincide, we conclude that the normalisation condition is equivalent to

$$\omega(\xi, \dots, \eta_i^{(m)}, \dots) = 1.$$

Henceforth a quasi-Abelian nilmanifold will be equipped with the right invariant volume form ω satisfying the normalisation condition above, and with the associated probability measure \mathcal{L} .

2.2.2. Quasi-Abelian nilflows. For any element $X \in \mathfrak{g}$, let $(\phi_X^t)_{t \in \mathbb{R}}$ denote the flow on M generated by X , that is, the flow given by right multiplication by the one-parameter subgroup $(\exp(tX))_{t \in \mathbb{R}}$:

$$(15) \quad \phi_X^t(\Gamma g) = \Gamma g \exp(tX), \quad \text{for all } \Gamma g \in M = \Gamma \backslash G.$$

Clearly this flow has an interesting dynamics only if $X \notin \mathfrak{a}$. Otherwise it is a linear flow on a toral fibre of the fibration in formula (12). It is a well-known classical result that the flow (ϕ_X^t) is ergodic, uniquely ergodic and minimal if and only if the projection of X in the Abelianised Lie algebra $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ is rationally independent of the lattice $\Gamma/[\Gamma, \Gamma]$.

Let $\alpha := (\alpha_i^{(m)}) \in \mathbb{R}^J$ and let

$$(16) \quad X_\alpha := \log \left[x^{-1} \exp \left(\sum_{(m,i) \in J} \alpha_i^{(m)} \tilde{\eta}_i^{(m)} \right) \right].$$

Return maps (to global transverse sections) of the flow generated by the above vector fields are readily computed as follows.

For $\theta \in \mathbb{T}^1$ let $\mathbb{T}_\theta^a = \text{pr}_2^{-1}(\{\theta\})$ denote the toral fibre above $\theta \in \mathbb{T}^1$ of the fibration pr_2 in formula (13).

For $m = 1, \dots, n$ let $\mathbb{Z}_\theta^{i_m}$ be the lattice defined as follows

$$(17) \quad \mathbf{s}^{(m)} \in \mathbb{Z}_\theta^{i_m} \Leftrightarrow \sum_{i=1}^i s_i^{(m)} \tilde{\eta}_i^{(m)} \in \mathfrak{a}^{(m)} \cap \text{Ad}(e^{-\theta\xi})(\Gamma).$$

If $\mathbb{Z}_\theta^a = \mathbb{Z}_\theta^{i_1} \times \dots \times \mathbb{Z}_\theta^{i_n}$ then we have

$$\mathbb{T}_\theta^a = \{\Gamma \exp(\mathbf{s}\tilde{\eta}) \exp(\theta\xi) \mid \mathbf{s} \in \mathbb{R}^a / \mathbb{Z}^a\} = \{\Gamma \exp(\theta\xi) \exp(\mathbf{s}\tilde{\eta}) \mid \mathbf{s} \in \mathbb{R}^a / \mathbb{Z}_\theta^a\}.$$

For all $m = 1, \dots, n$, let $\mathbb{T}_\theta^{i_m} \subset \mathbb{T}_\theta^a$ be the sub-torus defined as follows:

$$(18) \quad \begin{aligned} \mathbb{T}_\theta^{i_m} &= \{\Gamma \exp(\mathbf{s}^{(m)} \tilde{\eta}^{(m)}) \exp(\theta\xi) \mid \mathbf{s}^{(m)} \in \mathbb{R}^{i_m} / \mathbb{Z}^{i_m}\} \\ &= \{\Gamma \exp(\theta\xi) \exp(\mathbf{s}^{(m)} \tilde{\eta}^{(m)}) \mid \mathbf{s}^{(m)} \in \mathbb{R}^{i_m} / \mathbb{Z}_\theta^{i_m}\}. \end{aligned}$$

By construction, there exists a product decomposition

$$(19) \quad \mathbb{T}_\theta^a \approx \mathbb{T}_\theta^{i_1} \times \dots \times \mathbb{T}_\theta^{i_n}.$$

The torus \mathbb{T}_θ^a is a global section of the nilflow $(\phi_{X_\alpha}^t)_{t \in \mathbb{R}}$ on M , generated by $X_\alpha \in \mathfrak{g}$ (see formula (15)), and the product decomposition in formula (19) is $\{\phi_{X_\alpha}^t\}$ -invariant. The lemma below, which is classical (see [Fur81]), makes explicit its return maps.

Lemma 2.4. *The flow $(\phi_{X_\alpha}^t)_{t \in \mathbb{R}}$ on M is isomorphic to the suspension of its first return map $\Phi_{\alpha, \theta} : \mathbb{T}_\theta^a \rightarrow \mathbb{T}_\theta^a$, hence all return times are constant integer-valued functions on \mathbb{T}_θ^a . The first return map is a product*

$$\Phi_{\alpha, \theta} \approx \Phi_{\alpha^{(1)}, \theta} \times \dots \times \Phi_{\alpha^{(n)}, \theta} \quad \text{on} \quad \mathbb{T}_\theta^{i_1} \times \dots \times \mathbb{T}_\theta^{i_n},$$

and for every $m = 1, \dots, n$ the factor map $\Phi_{\alpha^{(m)}, \theta}$ is given in the coordinates $\mathbf{s}^{(m)} \in \mathbb{R}^{i_m} \bmod \mathbb{Z}_\theta^{i_m}$ by the formulas

$$(20) \quad \begin{aligned} \Phi_{\alpha^{(m)}, \theta}(\mathbf{s}^{(m)}) &= \Phi_{\alpha^{(m)}}(\mathbf{s}^{(m)}) = (s_1^{(m)} + \alpha_1^{(m)}, \dots, \\ &\quad s_j^{(m)} + s_{j-1}^{(m)} + \alpha_j^{(m)}, \dots, s_{i_m}^{(m)} + s_{i_m-1}^{(m)} + \alpha_{i_m}^{(m)}). \end{aligned}$$

For any $N \in \mathbb{Z}$, the N -th return map is a product

$$\Phi_{\alpha, \theta}^N \approx \Phi_{\alpha^{(1)}, \theta}^N \times \dots \times \Phi_{\alpha^{(n)}, \theta}^N \quad \text{on} \quad \mathbb{T}_\theta^{i_1} \times \dots \times \mathbb{T}_\theta^{i_n},$$

and for every $m = 1, \dots, n$ the factor map $\Phi_{\alpha^{(m)}, \theta}^N$ is given in the coordinates $\mathbf{s}^{(m)} \in \mathbb{R}^{i_m} \bmod \mathbb{Z}_\theta^{i_m}$ by the formulas

$$(21) \quad \begin{aligned} \Phi_{\alpha^{(m)}, \theta}^N(\mathbf{s}^{(m)}) &= (s_1^{(m)} + N \alpha_1^{(m)}, s_2^{(m)} + N(s_1^{(m)} + \alpha_2^{(m)}) + \binom{N}{2} \alpha_1^{(m)}, \\ &\quad \dots, s_{i_m}^{(m)} + \sum_{i=1}^{i_m-1} \binom{N}{i} (s_{i_m-i}^{(m)} + \alpha_{i_m-i+1}^{(m)}) + \binom{N}{i_m} \alpha_1^{(m)}). \end{aligned}$$

Proof. By construction, the factorisations in formula (14) and, more generally, in formula (19) are induced by the splitting of the vector space \mathfrak{a} into generalised eigenspaces of linear map $\text{ad}(X) : \mathfrak{a} \rightarrow \mathfrak{a}$. It follows that the all time- t maps of the flows $(\phi_{X_\alpha}^t)_{t \in \mathbb{R}}$ on \mathbb{T}_θ^a are isomorphic to the products of their restrictions to the tori $\mathbb{T}_\theta^{i_m} \subset \mathbb{T}_\theta^a$. For any given $m = 1, \dots, n$, let us compute the restriction of the time-1 map of the flow to the torus $\mathbb{T}_\theta^{i_m} \subset \mathbb{T}_\theta^a$.

Let $\Gamma \exp(\theta\xi) \exp(\sum_{j=1}^{i_m} s_j^{(m)} \tilde{\eta}_j^{(m)}) \in \mathbb{T}_\theta^{i_m}$. We have

$$\begin{aligned} \exp(\sum_{j=1}^{i_m} s_j^{(m)} \tilde{\eta}_j^{(m)}) \exp(X_\alpha) &= \exp(\sum_{j=1}^{i_m} s_j^{(m)} \tilde{\eta}_j^{(m)}) x^{-1} \exp(\sum_{j=1}^{i_m} \alpha_j^{(m)} \tilde{\eta}_j^{(m)}) \\ &= x^{-1} \exp[(s_1^{(m)} + \alpha_1^{(m)}) \tilde{\eta}_1^{(m)} + \sum_{j=1}^{i_m-1} (s_j^{(m)} + s_{j+1}^{(m)} + \alpha_{j+1}^{(m)}) \tilde{\eta}_{j+1}^{(m)}]. \end{aligned}$$

In fact, the following identity holds:

$$\begin{aligned} x \exp(\sum_{j=1}^{i_m} s_j^{(m)} \tilde{\eta}_j^{(m)}) x^{-1} &= \exp\left[e^{\text{ad}(\xi)} (\sum_{j=1}^{i_m} s_j^{(m)} \tilde{\eta}_j^{(m)})\right] \\ &= \exp[s_1^{(m)} \tilde{\eta}_1^{(m)} + \sum_{j=1}^{i_m-1} (s_j^{(m)} + s_{j+1}^{(m)}) \tilde{\eta}_{j+1}^{(m)}]. \end{aligned}$$

Since $x \in \Gamma$, it follows that

$$\begin{aligned} \Gamma \exp(\theta\xi) \exp(\sum_{j=1}^{i_m} s_j^{(m)} \tilde{\eta}_j^{(m)}) \exp(X_\alpha) \\ = \Gamma \exp(\theta\xi) \exp[(s_1^{(m)} + \alpha_1^{(m)}) \tilde{\eta}_1^{(m)} + \sum_{j=1}^{i_m-1} (s_j^{(m)} + s_{j+1}^{(m)} + \alpha_{j+1}^{(m)}) \tilde{\eta}_{j+1}^{(m)}]. \end{aligned}$$

The above formula implies that $t = 1$ is a return time of the restriction of the flow $(\phi_{X_\alpha}^t)_{t \in \mathbb{R}}$ to $\mathbb{T}_\theta^{i_m} \subset M$, for all $m = 1, \dots, n$, and the map (20) is the corresponding return map. In addition, $t = 1$ is the first return time, since it is the first return time of the projection onto $M^{(k)} \approx \mathbb{T}^{n+1}$ of the restriction of the flow $(\phi_{X_\alpha}^t)_{t \in \mathbb{R}}$ to the torus $\mathbb{T}^{i_m} \subset M$.

Finally, formula (21) for the N -th return map follows from formula (20) by induction on $N \in \mathbb{N}$. \square

2.3. Weyl sums as ergodic integrals. Let $P_k := P_k(N) \in \mathbb{R}[N]$ be a polynomial of degree $k \geq 2$:

$$P_k(N) := \sum_{j=0}^k a_j N^j.$$

A *Weyl sum* of degree $k \geq 2$ is the sum

$$W(P_k, f; N) = \sum_{\ell=0}^{N-1} f(P_k(\ell)),$$

for any $N \in \mathbb{N}$ and for any smooth periodic function $f \in C^\infty(\mathbb{T}^1)$. Classical (complete) Weyl sums are obtained as a particular case when the function f is the exponential function, that is,

$$f(s) = e(s) := \exp(2\pi i s), \quad s \in \mathbb{T}^1.$$

For any $(\alpha, \mathbf{s}) \in \mathbb{R}^k \times \mathbb{R}^k / \mathbb{Z}^k$, let $P_k(\alpha, \mathbf{s}, N) \in \mathbb{R}[N]$ be the polynomial of degree $k \geq 1$ defined (modulo \mathbb{Z}) as follows:

$$P_k(\alpha, \mathbf{s}, N) := \binom{N}{k} \alpha_1 + \sum_{j=1}^{k-1} \binom{N}{j} (s_{k-j} + \alpha_{k-j+1}) + s_k.$$

The following elementary result holds.

Lemma 2.5. *The map $(\alpha, \mathbf{s}) \rightarrow P_k(\alpha, \mathbf{s}, N)$ sends $\mathbb{R}^k \times \mathbb{R}^k / \mathbb{Z}^k$ onto the space $\mathbb{R}[N]$ of real polynomials (modulo \mathbb{Z}) of degree $k \geq 1$. The leading coefficient $a_k \in \mathbb{R}$ of the polynomial $P_k(\alpha, \mathbf{s}, N)$ is given by the formula:*

$$a_k = \frac{\alpha_1}{k!}.$$

More generally, the coefficient a_j of the the term of degree j of $P_k(\alpha, \mathbf{s}, N)$ is function

$$a_j = a_j(\alpha_1, \alpha_2 + s_1, \dots, \alpha_{k-j+1} + s_{k-j}), \quad \text{for } j = 1, \dots, k-1.$$

linear in each variable. For the constant term we have $a_0 = s_k$.

Let $\alpha \in \mathbb{R}^J$ and let $X_\alpha \in \mathfrak{g}$ be the vector field on M given by formula (16) and let B_α^T be the Birkhoff averaging operator defined, for all $f \in C^\infty(M)$ and all $(x, T) \in M \times \mathbb{R}$, by the formula

$$(22) \quad B_{X_\alpha}^T(x)(f) = \frac{1}{T} \int_0^T f \circ \phi_{X_\alpha}^t(x) dt.$$

The Weyl sums $\{W(P_k, f; N)\}_{N \in \mathbb{N}}$ for any smooth function $f \in C^\infty(\mathbb{T}^1)$ are special Birkhoff averages $B_\alpha^T(x)(F)$, hence it possible to derive bounds on Weyl sums from Sobolev bounds on the Birkhoff averaging operators B_α^T introduced above.

Definition 2.6. A quasi-Abelian Lie algebra on two generators (X, Y_1) is a *filiform* Lie algebra. A (generalised) Jordan basis $\{X, Y_1, \dots, Y_k\}$ of a quasi-Abelian filiform Lie algebra will be called a *(generalised) filiform basis*. A quasi-Abelian Lie group on two generators is a filiform Lie group. By definition, the Lie algebra of a quasi-Abelian filiform Lie group is a quasi-Abelian filiform Lie algebra. The quotient of a quasi-Abelian filiform group by a co-compact lattice is called a quasi-Abelian filiform nilmanifold.

Let $M = \Gamma \backslash G$ be a compact quasi-Abelian filiform nilmanifold. The torus \mathbb{T}_0^k denotes, as above, the set $(\Gamma \cap A) \backslash A \subset M$, that is, the orbit of the coset Γ under the action of the Abelian subgroup A . Let $\{\xi, \tilde{\eta}_1, \dots, \tilde{\eta}_k\}$ be the basis of the quasi-Abelian filiform Lie algebra \mathfrak{g} of G given (for the general quasi-Abelian case) in formulas (6), (7).

Let $\mathbf{s} = (s_1, \dots, s_k) \in \mathbb{R}^k / \mathbb{Z}^k$ denote the point $\Gamma \exp(s_1 \tilde{\eta}_1 + \dots + s_k \tilde{\eta}_k) \in \mathbb{T}_0^k$.

Lemma 2.7. *For any $\alpha \in \mathbb{R}^k$ there exist bounded injective linear operator*

$$F = F_\alpha : L^2(\mathbb{T}^1) \rightarrow L^2(M)$$

such that the following holds. For any $r \geq 0$, the operator F maps $H^r(\mathbb{T}^1)$ continuously into $H^r(M)$; moreover, for any $r > 1/2$, there exists a constant $C_r > 0$ such that, for any function $f \in H^r(\mathbb{T}^1)$, for all $(\mathbf{s}, N) \in \mathbb{T}_0^k \times \mathbb{N}$, we have

$$(23) \quad \left| \sum_{\ell=0}^N f(P_k(\alpha, \mathbf{s}, \ell)) - N B_{X_\alpha}^N(\mathbf{s})(F(f)) \right| \leq C_r \|f\|_{H^r(\mathbb{T}^1)}.$$

Proof. For any $\varepsilon \in]0, 1/2[$ the map

$$(24) \quad (\mathbf{s}, t) \in \mathbb{T}_0^k \times]-\varepsilon, \varepsilon[\mapsto \phi_{X_\alpha}^t(\mathbf{s}) = \mathbf{s} \exp(tX_\alpha)$$

is an embedding of $\mathbb{T}_0^k \times]-\varepsilon, \varepsilon[$ onto a tubular neighbourhood \mathcal{U}_ε of $\mathbb{T}_0^k \subset M$.

Let $pr : \mathbb{T}_0^k \rightarrow \mathbb{T}^1$ be the projection on the circle \mathbb{T}^1 defined as follows

$$pr(\mathbf{s}) = s_k, \quad \text{for all } \mathbf{s} \in \mathbb{T}_0^k.$$

Let $\chi \in C_0^\infty(]-\varepsilon, \varepsilon[)$ be any function such that $\int_{\mathbb{R}} \chi(\tau) d\tau = 1$ and let $C_1 = \|\chi\|_\infty$. For any $f \in L^2(\mathbb{T}^1)$, let $F(f) \in L^2(M)$ be the function defined on the open set \mathcal{U}_ε as

$$(25) \quad F(f)(\phi_{X_\alpha}^t(\mathbf{s})) = \chi(t)(f(pr(\mathbf{s}))), \quad (\mathbf{s}, t) \in \mathbb{T}_0^k \times]-\varepsilon, \varepsilon[.$$

We then extend the function $F(f)$ as zero on $M \setminus \mathcal{U}_\varepsilon$.

The function $F(f)$ is well-defined and square-integrable on M since χ is smooth and the map (24) is an embedding. Moreover, it follows from the definition (25) that $F(f) \in C^0(M)$ whenever $f \in C^0(\mathbb{T}^1)$ and $F(f) \in H^r(M)$ whenever $f \in H^r(\mathbb{T}^1)$, for any $r \geq 0$.

Let $f \in C^0(\mathbb{T}^1)$. We claim that, by the definition (25) of the function $F(f)$, for all $(\mathbf{s}, N) \in \mathbb{T}_0^k \times \mathbb{N}$, we have

$$(26) \quad \int_{-\varepsilon}^{N+\varepsilon} F(f) \circ \phi_{X_\alpha}^t(\mathbf{s}) dt = \sum_{\ell=0}^N f(P_k(\alpha, \mathbf{s}, \ell)).$$

In fact, let $\Phi_\alpha^\ell : \mathbb{T}_0^k \rightarrow \mathbb{T}_0^k$ be the ℓ -th return map of the flow $\{\phi_{X_\alpha}^t\}_{t \in \mathbb{R}}$. By Lemma 2.4 and by definition (25), for all $(\mathbf{s}, \ell) \in \mathbb{T}_0^k \times \mathbb{N}$,

$$pr \circ \Phi_\alpha^\ell(\mathbf{s}) = P_k(\alpha, \mathbf{s}, \ell),$$

hence, for all $f \in C^0(\mathbb{T}^1)$,

$$\begin{aligned} \int_{\ell-\varepsilon}^{\ell+1-\varepsilon} F(f) \circ \phi_{X_\alpha}^t(\mathbf{s}) dt &= \int_{\ell-\varepsilon}^{\ell+1-\varepsilon} F(f) \circ \phi_{X_\alpha}^t(x) dt \\ &= \int_{-\varepsilon}^{\varepsilon} F(f) \circ \phi_{X_\alpha}^\tau(\Phi_\alpha^\ell(\mathbf{s})) d\tau = \left(\int_{-\varepsilon}^{\varepsilon} \chi(\tau) d\tau \right) f(P_k(\alpha, \mathbf{s}, \ell)). \end{aligned}$$

The claim is therefore proved.

It follows by formula (26) that

$$\left| \int_0^N F(f) \circ \phi_{X_\alpha}^t(\mathbf{s}) dt - \sum_{\ell=0}^N f(P_k(\alpha, \mathbf{s}, \ell)) \right| \leq 2\varepsilon \|F(f)\|_\infty.$$

By the Sobolev embedding theorem $H^r(\mathbb{T}^1) \subset C^0(\mathbb{T}^1)$, for any $r > 1/2$, and there exists a constant $c_r > 0$ such that $\|f\|_\infty \leq c_r \|f\|_{H^r(\mathbb{T}^1)}$; since by definition $\|F(f)\|_\infty \leq \|\chi\|_\infty \|f\|_\infty$, the inequality (23) follows and the argument is concluded. \square

The problem of establishing bounds on Weyl sums is thus reduced to that of bounds for the nilpotent averages (22).

3. A SOBOLEV TRACE THEOREM

We prove below a Sobolev trace theorem for nilpotent orbits. According to this theorem, the uniform norm of an orbital (ergodic) integral is bounded in terms of the *average width* of the orbit segment times the *transverse Sobolev norms* of the function, with respect to a given basis of the Lie algebra. The average width of an orbit segment is a positive number which measures the weighted frequency of close returns of the orbit segment close to itself. Its definition is in fact very general

and our theorems can be generalised to the case of Lie groups with a codimension one ideal (Abelian or not). Bounds on the average width (for rescaled bases) of orbits segment of quasi-Abelian nilflows will be established in Section 5.

Let $M = \Gamma \backslash G$ be a quasi-Abelian nilmanifold and let ω be the associated volume form. We shall consider general ordered bases $\mathcal{F} := (X, Y)$ of \mathfrak{g} such that $Y := (Y_1, \dots, Y_a)$ is a basis of \mathfrak{a} .

Definition 3.1. An *adapted basis* of the Lie algebra \mathfrak{g} is an ordered basis $(X, Y) := (X, Y_1, \dots, Y_a)$ of \mathfrak{g} such that $X \notin \mathfrak{a}$ and $Y := (Y_1, \dots, Y_a)$ is a basis of the Abelian ideal $\mathfrak{a} \subset \mathfrak{g}$.

A *strongly adapted basis* $(X, Y) := (X, Y_1, \dots, Y_a)$ is an adapted basis such that the following holds:

- the system (X, Y_1, \dots, Y_n) is a system of generators of \mathfrak{g} , hence its projection is a basis of the Abelianisation $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ of the Lie algebra \mathfrak{g} ;
- the system (Y_{n+1}, \dots, Y_a) is a basis of the ideal $[\mathfrak{g}, \mathfrak{g}]$.

An adapted basis (X, Y_1, \dots, Y_a) is *normalised* if

$$\omega(X, Y_1, \dots, Y_a) = 1 \quad \text{on } M.$$

Note that, according to the above definition, all (generalised) Jordan bases (see Definition 2.2) are strongly adapted and normalised.

For any basis $Y = (Y_1, \dots, Y_a)$ of \mathfrak{a} and for all $\mathbf{s} := (s_1, \dots, s_a) \in \mathbb{R}^a$, it will be convenient to use the notations

$$\mathbf{s} \cdot Y := s_1 Y_1 + \dots + s_a Y_a.$$

3.1. Width of nilpotent orbits. Let $\mathcal{F} := (X, Y)$ be any normalised adapted basis of the quasi-Abelian nilpotent Lie algebra \mathfrak{g} . For any $x \in M$, let $\phi_x : \mathbb{R} \times \mathbb{R}^a \rightarrow M$ be the local embedding defined by

$$(27) \quad \phi_x(t, \mathbf{s}) = x \exp(tX) \exp(\mathbf{s} \cdot Y).$$

We omit the proofs of the following two elementary lemmata.

Lemma 3.2. For any $x \in M$ and any $f \in C^\infty(M)$ we have

$$\begin{aligned} \frac{\partial \phi_x^*(f)}{\partial t}(t, \mathbf{s}) &= \phi_x^*(Xf)(t, \mathbf{s}) + \sum_j s_j \phi_x^*([X, Y_j]f)(t, \mathbf{s}); \\ \frac{\partial \phi_x^*(f)}{\partial s_j} &= \phi_x^*(Y_j f), \quad \text{for all } j = 1, \dots, a. \end{aligned}$$

Lemma 3.3. For any $x \in M$, we have

$$\phi_x^*(\omega) = dt \wedge ds_1 \wedge \dots \wedge ds_a.$$

Let Leb_a denote the a -dimensional Lebesgue measure on \mathbb{R}^a .

Definition 3.4. For any open neighbourhood of the origin $O \subset \mathbb{R}^a$, let \mathcal{R}_O be the family of all a -dimensional symmetric (i.e. centred at the origin) rectangles $R \subset [-1/2, 1/2]^a$ such that $R \subset O$. The *inner width* of the open set $O \subset \mathbb{R}^a$ is the positive number

$$w(O) := \sup\{\text{Leb}_a(R) \mid R \in \mathcal{R}_O\}.$$

The *width function* of a set $\Omega \subset \mathbb{R} \times \mathbb{R}^a$ containing the line $\mathbb{R} \times \{0\}$ is the function $w_\Omega : \mathbb{R} \rightarrow [0, 1]$ defined as follows:

$$w_\Omega(\tau) := w(\{\mathbf{s} \in \mathbb{R}^a \mid (\tau, \mathbf{s}) \in \Omega\}), \quad \text{for all } \tau \in \mathbb{R}.$$

Definition 3.5. Let $\mathcal{F} = (X, Y)$ be any normalised adapted basis. For any $x \in M$ and $T > 1$, we consider the family $\mathcal{O}_{x,T}$ of all open sets $\Omega \subset \mathbb{R} \times \mathbb{R}^a$ satisfying:

- $[0, T] \times \{0\} \subset \Omega \subset \mathbb{R} \times [-1/2, 1/2]^a$;
- the map

$$\phi_x : \Omega \rightarrow M$$

defined by formula (27) is injective.

The *average width* of the orbit segment

$$\gamma_X(x, T) := \{x \exp(tX) \mid 0 \leq t \leq T\} = \{\phi_x(t, 0) \mid 0 \leq t \leq T\},$$

relative to the normalised adapted basis \mathcal{F} , is the positive real number

$$(28) \quad w_{\mathcal{F}}(x, T) := \sup_{\Omega \in \mathcal{O}_{x,T}} \left(\frac{1}{T} \int_0^T \frac{ds}{w_\Omega(s)} \right)^{-1}.$$

The *average width* of the nilmanifold M , relative to the normalised adapted basis \mathcal{F} , at a point $y \in M$ is the positive real number

$$(29) \quad w_{\mathcal{F}}(y) := \sup\{w_{\mathcal{F}}(x, 1) \mid y \in \gamma_X(x, 1)\}.$$

3.2. Sobolev a priori bounds. Let $\Delta_{\mathbf{s}}$ denote the Euclidean Laplace operator on \mathbb{R}^a :

$$\Delta_{\mathbf{s}} := - \sum_{j=1}^a \frac{\partial^2}{\partial s_j^2}.$$

For any $\sigma \in \mathbb{R}$, let $W^\sigma(R)$ denote the standard Sobolev space on a bounded open rectangle $R \subset \mathbb{R}^a$. The following lemma can be derived from the standard Sobolev embedding theorem for the unit cube $[-1/2, 1/2]^a$ using a rescaling argument.

Lemma 3.6. *For any $\sigma > a/2$, there exists a positive constant $C := C(a, \sigma)$ such that for any (symmetric) open rectangle $R \subset [-1/2, 1/2]^a$ and for any function $f \in W^\sigma(R)$, we have*

$$|f(0)| \leq \frac{C}{\text{Vol}(R)^{1/2}} \left(\int_R |(I + \Delta_{\mathbf{s}})^{\sigma/2} f(\mathbf{s})|^2 d^a \mathbf{s} \right)^{1/2}.$$

Lemma 3.7. *Let $I \subset \mathbb{R}$ be a bounded interval and let $\Omega \subset \mathbb{R} \times \mathbb{R}^a$ be a Borel set containing the segment $I \times \{0\} \subset \mathbb{R} \times \mathbb{R}^a$. For any $\sigma > a/2$, there exists a positive constant $C_{a,\sigma}$ such that, for all functions $F \in C^\infty(\Omega)$ and for all $t \in I$, we have*

$$(30) \quad \left| \int_I F(t, 0) dt \right|^2 \leq C_{a,\sigma}^2 \left(\int_I \frac{d\tau}{w_\Omega(\tau)} \right) \int_\Omega |(I + \Delta_{\mathbf{s}})^{\frac{\sigma}{2}} F(\tau, \mathbf{s})|^2 d\tau d^a \mathbf{s};$$

and

$$(31) \quad |F(t, 0)| \leq \frac{C_{a,\sigma}}{|I|} \left(\int_I \frac{d\tau}{w_\Omega(\tau)} \right)^{1/2} \left[\left(\int_\Omega |(I + \Delta_{\mathbf{s}})^{\frac{\sigma}{2}} F(\tau, \mathbf{s})|^2 d\tau d^a \mathbf{s} \right)^{1/2} + |I| \left(\int_\Omega |(I + \Delta_{\mathbf{s}})^{\frac{\sigma}{2}} \partial_t F(\tau, \mathbf{s})|^2 d\tau d^a \mathbf{s} \right)^{1/2} \right].$$

Proof. For any $t \in \mathbb{R}$, let $\Omega_t := \{\mathbf{s} \in \mathbb{R}^a \mid (t, \mathbf{s}) \in \Omega\}$. By the definition of the width function (see Definition 3.4) and by the standard Sobolev embedding theorem for bounded rectangles in \mathbb{R}^a , it follows that there exists a constant $C_{a,\sigma} > 0$ such that, for any function $G \in C^\infty(\Omega)$ and for any $\tau \in I$,

$$|G(\tau, 0)| \leq \frac{C_{a,\sigma}}{w_\Omega(\tau)^{1/2}} \left(\int_{\Omega_\tau} |(I + \Delta_{\mathbf{s}})^{\frac{\sigma}{2}} G(\tau, \mathbf{s})|^2 d^a \mathbf{s} \right)^{1/2}.$$

Then by Hölder's inequality it follows that

$$(32) \quad \left(\int_I |G(\tau, 0)| d\tau \right)^2 \leq C_{a,\sigma}^2 \left(\int_I \frac{d\tau}{w_\Omega(\tau)} \right) \int_\Omega |(I + \Delta_{\mathbf{s}})^{\frac{\sigma}{2}} G(\tau, \mathbf{s})|^2 d^a \mathbf{s} d\tau.$$

Taking $G = F$ in the above formula yields the estimate in formula (30). To see (31) observe that, by the fundamental theorem of calculus and the mean value theorem, for any $t \in \mathbb{R}$ there exists $t_0 = t_0(t) \in I$ such that

$$F(t, 0) = \frac{1}{|I|} \int_I F(\tau, 0) d\tau + \int_{t_0}^t \partial_t F(\tau, 0) d\tau,$$

which implies

$$|F(t, 0)| \leq \int_I \left| \frac{F(\tau, 0)}{|I|} \right| d\tau + \int_I |\partial_t F(\tau, 0)| d\tau.$$

The estimate in formula (31) then follows by applying the bound in formula (32) to the functions $G = F/|I|$ and $G = \partial_t F$. The statement is proved. \square

Definition 3.8. Given any normalised adapted basis $\mathcal{F} = (X, Y)$, let Δ_Y be the second order differential operator defined as follows:

$$(33) \quad \Delta_Y := - \sum_{j=1}^a Y_j^2.$$

For any $\sigma \geq 0$, let $|\cdot|_{\mathcal{F},\sigma}$ be the *transverse Sobolev norm* defined as follows: for all functions $f \in C^\infty(M)$, let

$$(34) \quad |f|_{\mathcal{F},\sigma} := \|(I + \Delta_Y)^{\frac{\sigma}{2}} f\|_{L^2(M)}.$$

The completion of $C^\infty(M)$ with respect to the norm $|\cdot|_{\mathcal{F},\sigma}$ is denoted $W^\sigma(M, \mathcal{F})$. Endowed with this norm $W^\sigma(M, \mathcal{F})$ is a Hilbert space.

The following version of the Sobolev embedding theorem holds.

Theorem 3.9. *Let $\mathcal{F} = (X, Y)$ be any normalised adapted basis. For any $\sigma > a/2$, there exist positive constants $C_{a,\sigma}$, C_σ such that, for all functions $u \in W^{\sigma+1}(M, \mathcal{F})$ such that $Xu \in W^\sigma(M, \mathcal{F})$ and for all $y \in M$, we have*

$$|u(y)| \leq \frac{C_{a,\sigma}}{w_{\mathcal{F}}(y)^{1/2}} \left\{ |u|_{\mathcal{F},\sigma} + |Xu|_{\mathcal{F},\sigma} + C_\sigma \sum_{j=1}^a |[X, Y_j]u|_{\mathcal{F},\sigma} \right\}.$$

Proof. Let $y \in M$ be a given point and let $x \in M$ be any point such that $y \in \gamma_X(x, 1)$. Let $\Omega \subset \mathbb{R} \times [-1/2, 1/2]^a$ be an open set containing $[0, 1] \times \{0\}$ such that the map ϕ_x defined by (27) is injective on Ω . Let $F(t, \mathbf{s}) = u \circ \phi_x(t, \mathbf{s})$ for all $(t, \mathbf{s}) \in \Omega$. By Lemma 3.2 we have

$$(35) \quad \begin{aligned} \partial_t F(t, \mathbf{s}) &= (Xu + \sum_j s_j [X, Y_j]u) \circ \phi_x, \\ \Delta_{\mathbf{s}} F &= (\Delta_Y u) \circ \phi_x. \end{aligned}$$

By Lemma 3.3 and by the fact that the basis \mathcal{F} is normalised, it follows that the map ϕ_x maps the measure $d\tau d^a \mathbf{s}$ to the measure \mathcal{L} ; thus

$$\begin{aligned} \left(\int_{\Omega} |(I + \Delta_{\mathbf{s}})^{\frac{\sigma}{2}} F(\tau, \mathbf{s})|^2 d\tau d^a \mathbf{s} \right)^{1/2} &= \left(\int_{\phi_x(\Omega)} |(I + \Delta_Y)^{\frac{\sigma}{2}} u|^2 d\mathcal{L} \right)^{1/2} \\ &\leq \|(I + \Delta_Y)^{\frac{\sigma}{2}} u\|_{L^2(M)}. \end{aligned}$$

For all $\tau \in I$, let $\Omega_{\tau} := \{\mathbf{s} \in \mathbb{R}^a | (\tau, \mathbf{s}) \in \Omega\}$. By a direct computation for $\sigma \in \mathbb{N}$ and by the interpolation property of Sobolev norms in the general case, there exists a constant $C_{\sigma} > 0$ such that, for all $\tau \in I$, for all $j \in \{1, \dots, a\}$ and for all $f \in W^{\sigma}(\Omega_{\tau})$, we have

$$\int_{\Omega_{\tau}} |(I + \Delta_{\mathbf{s}})^{\frac{\sigma}{2}} (s_j f)|^2 d^a \mathbf{s} \leq C_{\sigma}^2 \int_{\Omega_{\tau}} |(I + \Delta_{\mathbf{s}})^{\frac{\sigma}{2}} f|^2 d^a \mathbf{s}.$$

By the above estimates it then follows that

$$\begin{aligned} &\left(\int_{\Omega} |(I + \Delta_{\mathbf{s}})^{\frac{\sigma}{2}} \partial_t F(\tau, \mathbf{s})|^2 d\tau d^a \mathbf{s} \right)^{1/2} \\ &= \left(\int_{\Omega} \left| (I + \Delta_{\mathbf{s}})^{\frac{\sigma}{2}} (Xu + \sum_j s_j [X, Y_j]u) \circ \phi_x(\tau, \mathbf{s}) \right|^2 d\tau d^a \mathbf{s} \right)^{1/2} \\ &\leq \|(I + \Delta_Y)^{\frac{\sigma}{2}} Xu\|_{L^2(M)} + C_{\sigma} \sum_{j=1}^a \|(I + \Delta_Y)^{\frac{\sigma}{2}} [X, Y_j]u\|_{L^2(M)}. \end{aligned}$$

By applying the estimate (31) of Lemma 3.7 to the function $\phi_x^* u$, we obtain

$$\begin{aligned} |u(y)| \leq C_{a,\sigma} \left(\int_0^1 \frac{d\tau}{w_{\Omega}(\tau)} \right)^{1/2} &\left[\|(I + \Delta_Y)^{\frac{\sigma}{2}} u\|_{L^2(M)} \right. \\ &+ \|(I + \Delta_Y)^{\frac{\sigma}{2}} Xu\|_{L^2(M)} \\ &\left. + C_{\sigma} \sum_{j=1}^a \|(I + \Delta_Y)^{\frac{\sigma}{2}} [X, Y_j]u\|_{L^2(M)} \right]. \end{aligned}$$

Since the above inequality holds for all open sets $\Omega \subset \mathbb{R} \times [-1/2, 1/2]^a$ containing $[0, 1] \times \{0\}$, for which the restriction to Ω of the map ϕ_x is injective, we can replace the term $\left(\int_0^1 d\tau / w_{\Omega}(\tau) \right)^{1/2}$ in the above inequality by its infimum over all such sets, that is, by the lower bound $1/w_{\mathcal{F}}(x, 1)^{1/2}$ (see formula (28) of Definition 3.5). Finally, the statement follows by taking the infimum over all points $x \in M$ such that $y \in \gamma_X(x, 1)$ (see formula (29) of Definition 3.5). \square

3.3. Nilpotent averages. For a vector field X on M and $x \in M$ the Birkhoff ergodic average $B_X^T(x)$ is defined as follows: for all $f \in L^2(M)$,

$$B_X^T(x)(f) := \frac{1}{T} \int_0^T f \circ \phi_X^t(x) dt, \quad \text{for all } T \in \mathbb{R}_+.$$

where ϕ_X^t is the flow generated by the vector field X . The following Sobolev estimates for the linear functional $B_X^T(x)$ holds on $W^{\sigma}(M, \mathcal{F})$.

Theorem 3.10. *Let $\mathcal{F} = (X, Y)$ be any normalised adapted basis. For any $\sigma > a/2$, there exists a positive constant $C_{a,\sigma}$ such that, for all functions $f \in W^{\sigma}(M, \mathcal{F})$, for*

all $T \in [1, +\infty)$ and all $x \in M$ we have

$$|B_X^T(x)(f)| \leq \frac{C_{a,\sigma}}{T^{1/2}w_{\mathcal{F}}(x,T)^{1/2}}|f|_{\mathcal{F},\sigma}.$$

Proof. Let $\Omega \subset \mathbb{R} \times [-1/2, 1/2]^a$ be any open set containing $[0, T] \times \{0\}$ for which the restriction to Ω of the map ϕ_x is injective (that is, let $\Omega \in \mathcal{O}_{x,T}$). The estimate (30) of Lemma 3.7 applied to the function $F = f \circ \phi_x$ yields

$$\begin{aligned} |B_X^T(x)(f)| &= \left| \frac{1}{T} \int_0^T F(t, 0) dt \right| \\ &\leq C_{a,\sigma} \frac{1}{T} \left(\int_I \frac{d\tau}{w_{\Omega}(\tau)} \right)^{1/2} \left(\int_{\Omega} |(I + \Delta_{\mathbf{s}})^{\frac{\sigma}{2}} F(\tau, \mathbf{s})|^2 d\tau d^a \mathbf{s} \right)^{1/2} \\ &\leq C_{a,\sigma} \frac{1}{T} \left(\int_I \frac{d\tau}{w_{\Omega}(\tau)} \right)^{1/2} \|(I + \Delta_Y)^{\frac{\sigma}{2}} f\|_{L^2(M)}. \end{aligned}$$

As this inequality holds true for all $\Omega \in \mathcal{O}_{x,T}$, we can replace the term $\left(\int_0^T d\tau / w_{\Omega}(\tau) \right)^{1/2}$ in the above inequality by its infimum over all such sets, i.e. the lower bound $T^{1/2}/w_{\mathcal{F}}(x,T)^{1/2}$. \square

4. THE COHOMOLOGICAL EQUATION

In this section we prove a priori Sobolev estimates on the Green operator for the cohomological equation $Xu = f$ of a quasi-Abelian nilflow with generator $X \in \mathfrak{g} \setminus \mathfrak{a}$ and on the distributional obstructions to existence of solutions (that is, on invariant distributions). We then derive bounds on Sobolev norms the Green operator and on the scaling of invariant distributions under a group of dilations of the quasi-Abelian Lie algebra. We recall that this analysis is motivated, on the one hand, by the well-known elementary fact that ergodic integrals of coboundaries with bounded transfer function (that is, of all derivatives of bounded functions along the flow) are uniformly bounded, on the other hand, by the heuristic principle that the growth of ergodic integrals is related to the scaling of the invariant distributions under an appropriate renormalisation group action.

4.1. Irreducible unitary representations.

4.1.1. Representation models. Kirillov's theory yields the following complete classification of irreducible unitary representations of filiform Lie groups (up to unitary equivalence).

Let \mathfrak{a}^* be the space of \mathbb{R} -linear forms on \mathfrak{a} . For any $\Lambda \in \mathfrak{a}^*$ denote by $\exp \iota \Lambda$ the character χ_{Λ} of A defined by $\chi_{\Lambda}(g) := \exp(\iota \Lambda(Y))$, for $g = \exp Y$ with $Y \in \mathfrak{a}$.

The infinite dimensional irreducible representations of G are unitarily equivalent to the representations $\text{Ind}_A^G(\Lambda)$, obtained by inducing from A to G a character $\chi = \exp \iota \Lambda$ not vanishing on $[\mathfrak{a}, \mathfrak{a}]$. In addition, two linear forms Λ and Λ' determine unitarily equivalent representations if and only if they belong to the same co-ajoint orbit.

Restricting the function of $\text{Ind}_A^G(\Lambda)$ to the subgroup $\exp(tX)$, $t \in \mathbb{R}$, yields the following models for the unitary representations $\text{Ind}_A^G(\Lambda)$.

For $X \in \mathfrak{g} \setminus \mathfrak{a}$, $Y \in \mathfrak{a}$ and $\Lambda \in \mathfrak{a}^*$, we denote by $P_{\Lambda,Y}$ the polynomial function $x \rightarrow \Lambda(\text{Ad}(e^{xX})Y)$. Let π_{Λ}^X be the unitary representation of the quasi-Abelian

k -step nilpotent Lie group G on the Hilbert space $L^2(\mathbb{R})$ uniquely determined by the derived representation $D\pi_\Lambda^X$ of the filiform Lie algebra \mathfrak{g} given by the following formulas:

$$(36) \quad D\pi_\Lambda^X : \begin{cases} X \mapsto \frac{d}{dx} \\ Y \mapsto \iota P_{\Lambda,Y}(x) \end{cases} \quad \text{for all } Y \in \mathfrak{a}.$$

For each $\Lambda \in \mathfrak{a}^*$, not vanishing on $[\mathfrak{g}, \mathfrak{g}]$, the unitary representation π_Λ^X is irreducible and, by Kirillov's theory, each irreducible unitary representation of the quasi-Abelian k -step nilpotent Lie group G , which does not factor through a unitary representation of the Abelian quotient $G/[G, G]$, is unitarily equivalent to a representation of the form π_Λ^X described above.

Let

$$\mathfrak{a}_0^* = \{\Lambda \in \mathfrak{a}^* \mid \Lambda([\mathfrak{g}, \mathfrak{g}]) \neq 0\}.$$

Definition 4.1. For any $Y \in \mathfrak{a}$ we define its degree $d_Y \in \mathbb{N}$ with respect to the representation π_Λ^X to be the degree of the polynomial $P_{\Lambda,Y}$. For any adapted basis $\mathcal{F} = (X, Y)$ of the Lie algebra \mathfrak{g} let $(d_1, \dots, d_a) \in \mathbb{N}^a$ denote the degrees of the elements (Y_1, \dots, Y_a) . The degree of the representation π_Λ is then defined as the maximum of the degrees of the elements of any adapted basis.

Observe that the condition $\Lambda \in \mathfrak{a}_0^*$ is equivalent to $(d_1, \dots, d_a) \neq 0$.

For all $i = 1, \dots, a$ and $j = 1, \dots, d_i$, we let

$$(37) \quad \Lambda_i^{(j)}(\mathcal{F}) = (\Lambda \circ \text{ad}^j(X))(Y_i).$$

Then the representation π_Λ^X can be written as follows:

$$(38) \quad D\pi_\Lambda^X : \begin{cases} X \mapsto \frac{d}{dx} \\ Y_i \mapsto \iota \sum_{j=0}^{d_i} \frac{\Lambda_i^{(j)}(\mathcal{F})}{j!} x^j. \end{cases}$$

For any linear form $\Lambda \in \mathfrak{a}_0^*$, let $\mathfrak{I}_\Lambda \subset \mathfrak{a}$ be the subset defined as follows:

$$(39) \quad \mathfrak{I}_\Lambda := \bigcap_{i=0}^{k-1} \ker(\Lambda \circ \text{ad}^i(X)).$$

Since \mathfrak{g} is quasi-Abelian, the set $\mathfrak{I}_\Lambda \subset \mathfrak{g}$ is an ideal of the Lie algebra \mathfrak{g} . Let $G_\Lambda \subset G$ the normal subgroup defined by exponentiation of the ideal \mathfrak{I}_Λ . It is clear from the above definition that the ideal \mathfrak{I}_Λ , hence the subgroup G_Λ , depends only on the co-adjoint orbit of the form $\Lambda \in \mathfrak{a}^*$.

Lemma 4.2. *The irreducible unitary representation π_Λ^X of the quasi-Abelian Lie group G factors through a representation of the filiform Lie group G/G_Λ . In fact, for any adapted basis $\mathcal{F} := (X, Y)$ of the Lie algebra and for any element $Y_* \in \mathcal{F} \cap \mathfrak{a}$ of maximal degree $d \geq 1$ for the representation, the system*

$$(Y'_1, \dots, Y'_{d+1}) = (Y_*, \text{ad}(X)(Y_*), \dots, \text{ad}^d(X)(Y_*))$$

can be extended to an adapted basis $\mathcal{F}'_\Lambda := (X, Y'_1, \dots, Y'_a)$ with

$$\pi_\Lambda^X(Y'_i) = 0, \quad \text{for all } i \in \{d+2, \dots, a\}.$$

The basis (X, Y') is strongly adapted or Jordan if the basis (X, Y) is respectively strongly adapted or Jordan. In addition the coefficients of the change of basis matrix

$(C_{ij}^{Y,Y'}) \in M_a(\mathbb{R})$, that is, of the matrix such that

$$(Y'_1, \dots, Y'_a) = (Y_1, \dots, Y_a) \cdot C^{Y,Y'},$$

can be estimated as follows. Let

$$C_{\mathcal{F},\Lambda} := \Lambda(Y'_{d+1})^{-1} \max_{1 \leq i \leq a, 1 \leq j \leq d_i} |\Lambda_i^{(j)}(\mathcal{F})|.$$

There exists a constant $K_{\mathcal{F}} \geq 1$ (equal to one if the basis \mathcal{F} is Jordan) such that the following upper bound holds:

$$(40) \quad |C_{i,j}^{Y,Y'}| \leq K_{\mathcal{F}} C_{\mathcal{F},\Lambda} (1 + C_{\mathcal{F},\Lambda})^d, \quad \text{for all } i, j \in \{1, \dots, a\}.$$

Proof. Let $Y_* \in \{Y_1, \dots, Y_a\}$ be an element of maximal degree $d \geq 1$ and let

$$(Y'_1, \dots, Y'_{d+1}) = (Y_*, \text{ad}(X)(Y_*), \dots, \text{ad}^d(X)(Y_*)).$$

By construction, it follows that $\Lambda(Y'_i) \neq 0$ for all $i = 1, \dots, d+1$. Since the representation π_{Λ}^X has degree d , the ideal $\mathfrak{J}_{\Lambda} \subset \mathfrak{g}$ defined in formula (39) has at most codimension $d+1$. Hence

$$\mathfrak{a} = \bigoplus_{i=1}^{d+1} \mathbb{R} Y'_i \bigoplus \mathfrak{J}_{\Lambda}.$$

The representation π_{Λ}^X factorises through a representation of the filiform Lie group G_{Λ} of Lie algebra $\mathfrak{g}/\mathfrak{J}_{\Lambda}$. In fact, by construction and by the definition of induced representation

$$\pi_{\Lambda}^X(Y) = 0, \quad \text{for all } Y \in \mathfrak{J}_{\Lambda}.$$

For every $i \in \{1, \dots, a\}$, let $d_i \in \mathbb{N}$ denote the degree of the vector field $Y_i \in Y$. Let us consider the the following system of linear equations :

$$(41) \quad \Lambda(\text{ad}^{\ell}(X)(Y_i)) = \sum_{j=\ell}^{d_i} c_j^{(i)} \Lambda(Y'_{d-j+\ell+1}), \quad \forall \ell = 0, \dots, d_i, \forall i = d+2, \dots, a.$$

We claim that this system has a unique solution $(c_j^{(i)})$, $(i = d+2, \dots, a, j = 0, \dots, d_i)$ which satisfies the upper bounds

$$(42) \quad |c_j^{(i)}| \leq C_{\mathcal{F},\Lambda} (1 + C_{\mathcal{F},\Lambda})^{d_i-j}.$$

The proof of this claim proceeds by induction. For $\ell = d_i$ we have

$$\Lambda(\text{ad}^{d_i}(X)(Y_i)) = c_{d_i}^{(i)} \Lambda(Y'_{d+1}).$$

By the above equation, since $\Lambda(Y'_{d+1}) \neq 0$, the coefficients $c_{d_i}^{(i)} \in \mathbb{R}$ are uniquely defined for all $i = d+2, \dots, a$ and the following upper bound holds by definition:

$$|c_{d_i}^{(i)}| \leq |\Lambda(Y'_{d+1})|^{-1} |\Lambda(\text{ad}^{d_i}(X)(Y_i))| \leq C_{\mathcal{F},\Lambda}.$$

Let us assume the induction hypothesis that for $d_i \geq j > \ell$ the coefficients $c_j^{(i)} \in \mathbb{R}$ are uniquely defined and satisfy the upper bounds in formula (42). The coefficients $c_{\ell}^{(i)} \in \mathbb{R}$ can then be found for all $i = d+2, \dots, a$ by formula (41), which also implies that the following estimates hold:

$$|c_{\ell}^{(i)}| \leq |\Lambda(Y'_{d+1})|^{-1} (|\Lambda(\text{ad}^{\ell}(X)(Y_i))| + \sum_{j=\ell+1}^{d_i} |c_j^{(i)}| \cdot |\Lambda(Y'_{d-j+\ell+1})|).$$

Since by definition we have the bound

$$|\Lambda(Y'_j)| \leq C_{\mathcal{F},\Lambda} \cdot \Lambda(Y'_{d+1}), \quad \forall j = 1, \dots, d+1,$$

by the induction hypothesis we conclude that

$$|c_\ell^{(i)}| \leq C_{\mathcal{F},\Lambda} \left(1 + \sum_{j=\ell+1}^{d_i} C_{\mathcal{F},\Lambda} (1 + C_{\mathcal{F},\Lambda})^{d_i-j} \right) = C_{\mathcal{F},\Lambda} (1 + C_{\mathcal{F},\Lambda})^{d_i-\ell}.$$

We have therefore proved that the system in formula (41) has a unique solution which satisfies the estimates in formula (42) hold.

We now complete the system $\{Y'_1, \dots, Y'_{d+1}\}$ to obtain a basis of the Abelian Lie sub-algebra \mathfrak{a} .

Up to reordering the elements of the basis $Y \subset \mathfrak{a}$, it is not restrictive to assume that $\{Y'_1, \dots, Y'_{d+1}, Y_{d+2}, \dots, Y_a\}$ is a basis of \mathfrak{a} . For $i = d+2, \dots, a$, let

$$(43) \quad Y'_i := Y_i - \sum_{j=0}^{d_i} c_j^{(i)} Y'_{d-j+1}.$$

It follows from formula (43) that $\{Y'_1, \dots, Y'_{d+1}, Y'_{d+2}, \dots, Y'_a\}$ is a basis of \mathfrak{a} and from formulas (41) and (43) that $\{Y'_{d+2}, \dots, Y'_a\} \subset \mathfrak{I}_\Lambda$. If the basis (X, Y) is adapted, so is the the basis $(X, Y') := (X, Y'_1, \dots, Y'_a)$ since the systems $\{Y'_1, \dots, Y'_a\}$ and $\{Y_1, \dots, Y_a\}$ span the same subspace. If (X, Y) is strongly adapted, then up to reordering the elements of the basis $Y' \subset \mathfrak{a}$ the basis (X, Y') is also strongly adapted. In fact, by definition $Y'_2, \dots, Y'_{d+1} \in [\mathfrak{g}, \mathfrak{g}]$, hence the element $Y'_i \in [\mathfrak{g}, \mathfrak{g}]$ whenever $Y_i \in [\mathfrak{g}, \mathfrak{g}]$, for all $i \in \{d+2, \dots, a\}$. It follows from formulas (41) and (43) that if the basis $\mathcal{F} = (X, Y)$ is Jordan, then, up to reordering the elements of the basis $Y' \subset \mathfrak{a}$, the basis (X, Y') is Jordan as well.

The estimates in formula (40) can be derived from the upper bounds in formula (42) by formula (43). The constant $K_{\mathcal{F}} \geq 1$ is defined as follows. Let (a_{ij}) denote the matrix of the coordinates of the vectors Y'_1, \dots, Y'_{d+1} with respect to the basis $Y := \{Y_1, \dots, Y_a\} \subset \mathfrak{a}$ and let χ denote the indicator function of the set $\{d+2, \dots, a\}$. Let us define the constant

$$K_{\mathcal{F}} := \max_{1 \leq j \leq a} (\chi_j + \sum_{i=1}^{d+1} |a_{ij}|).$$

The estimates in formula (40) then follow from the estimates in formula (42) by the above definition and by formula (43). In the special case that the basis \mathcal{F} is Jordan, the above constant $K_{\mathcal{F}} = 1$ since by construction the set $\{Y'_1, \dots, Y'_{d+1}\}$ is a subset of $\{Y_1, \dots, Y_a\}$, disjoint from the subset $\{Y_{d+2}, \dots, Y_a\}$. \square

Motivated by the above lemma we introduce the following

Definition 4.3. A *generalised filiform basis* for an induced irreducible unitary representation π_Λ^X of degree $d \geq 1$ of a quasi-Abelian nilpotent Lie group is an adapted basis (X, Y_1, \dots, Y_a) such that (X, Y_1, \dots, Y_{d+1}) is a filiform basis for the generated filiform sub-algebra and $\{Y_{d+2}, \dots, Y_a\}$ is a basis of $\ker(\pi_\Lambda^X)$, that is,

$$[X, Y_i] = Y_{i+1}, \text{ for } 1 \leq i \leq d, \quad \text{and} \quad \pi_\Lambda^X(Y_i) = 0, \text{ for } d+2 \leq i \leq a.$$

According to Lemma 4.2, generalised filiform bases exist for all irreducible unitary representations π_Λ^X of non-zero degree of quasi-Abelian Lie groups and their norm can be bounded in terms of the linear functional $\Lambda \in \mathfrak{a}_0^*$.

4.1.2. *Sobolev norms.* We denote by $C^\infty(\pi_\Lambda^X)$ the space of C^∞ vectors of the irreducible unitary representation π_Λ^X defined by the formulas (36) and (38).

The transverse Sobolev norms introduced in formula (34) can be written in representation as follows. For the representation π_Λ^X , the transverse Laplace operator Δ_Y , introduced in formula (33), is represented as the operator of multiplication by the non-negative polynomial function

$$(44) \quad \Delta_{\Lambda, \mathcal{F}}(x) := \sum_{i=1}^a |P_{\Lambda, Y_i}(x)|^2 = \sum_{i=1}^a \left| \sum_{j=0}^{d_i} \frac{\Lambda_i^{(j)}(\mathcal{F})}{j!} x^j \right|^2.$$

Thus, the transverse Sobolev norms can be written as follows: for every $\sigma \geq 0$ and for every $f \in C^\infty(\pi_\Lambda^X)$,

$$|f|_{\mathcal{F}, \sigma} := \left(\int_{\mathbb{R}} [1 + \Delta_{\Lambda, \mathcal{F}}(x)]^{\frac{\sigma}{2}} |f(x)|^2 dx \right)^{1/2}.$$

4.2. **A priori estimates.** The unique distributional obstruction to the existence of solutions of the cohomological equation

$$(45) \quad Xu = f$$

in a given irreducible unitary representation π_Λ^X is the normalised X -invariant distribution $\mathcal{D}_\Lambda^X \in \mathcal{D}'(\pi_\Lambda^X)$ which can be written as

$$(46) \quad \mathcal{D}_\Lambda^X(f) := \int_{\mathbb{R}} f(x) dx, \quad \text{for all } f \in C^\infty(\pi_\Lambda^X).$$

The formal Green operator G_Λ^X for the the cohomological equation (45) is given by the formula

$$(47) \quad G_\Lambda^X(f)(x) := \int_{-\infty}^x f(y) dy, \quad \text{for all } f \in C^\infty(\pi_\Lambda^X).$$

It is not difficult to prove that the Green operator is well-defined on the kernel $\mathcal{K}^\infty(\pi_\Lambda^X)$ of the distribution \mathcal{D}_Λ^X on $C^\infty(\pi_\Lambda^X)$: for all $f \in \mathcal{K}^\infty(\pi_\Lambda^X)$, the function $G_\Lambda^X(f) \in C^\infty(\pi_\Lambda)$ and the following identities hold:

$$(48) \quad G_\Lambda^X(f)(x) := \int_{-\infty}^x f(y) dy = - \int_x^{+\infty} f(y) dy.$$

We prove below bounds on the transverse Sobolev norms $\|G_\Lambda^X(f)\|_{\tau, \mathcal{F}}$ for all functions $f \in \mathcal{K}^\infty(\pi_\Lambda^X)$.

For any $\sigma, \tau \in \mathbb{R}_+$ let

$$(49) \quad \begin{aligned} I_\sigma(\Lambda, \mathcal{F}) &:= \left(\int_{\mathbb{R}} \frac{dx}{[1 + \Delta_{\Lambda, \mathcal{F}}(x)]^\sigma} \right)^{1/2}; \\ J_\sigma^\tau(\Lambda, \mathcal{F}) &:= \left(\iint_{|y| \geq |x|} \frac{[1 + \Delta_{\Lambda, \mathcal{F}}(x)]^\tau}{[1 + \Delta_{\Lambda, \mathcal{F}}(y)]^\sigma} dx dy \right)^{1/2}. \end{aligned}$$

Lemma 4.4. *Let $\mathcal{D}_\Lambda^X \in \mathcal{D}'(\pi_\Lambda^X)$ be the distribution defined in formula (46). For any $\sigma \in \mathbb{R}_+$, the following holds:*

$$(50) \quad |\mathcal{D}_\Lambda^X|_{\mathcal{F}, -\sigma} := \sup_{f \neq 0} \frac{|\mathcal{D}_\Lambda^X(f)|}{|f|_{\mathcal{F}, \sigma}} = I_\sigma(\Lambda, \mathcal{F}).$$

Proof. It follows from the definitions by Hölder inequality. In fact,

$$(51) \quad \mathcal{D}_\Lambda^X(f) = \langle (1 + \Delta_{\Lambda, \mathcal{F}})^{-\frac{\sigma}{2}}, (1 + \Delta_{\Lambda, \mathcal{F}})^{\frac{\sigma}{2}} f \rangle_{L^2(\mathbb{R})}.$$

Since $|f|_{\mathcal{F}, \sigma} = |(1 + \Delta_{\Lambda, \mathcal{F}})^{\frac{\sigma}{2}} f|_0$, it follows that

$$\sup_{f \neq 0} \frac{|\mathcal{D}_\Lambda^X(f)|}{|f|_{\mathcal{F}, \sigma}} = |(1 + \Delta_{\Lambda, \mathcal{F}})^{-\frac{\sigma}{2}}|_{L^2(\mathbb{R})} = I_\sigma(\Lambda, \mathcal{F}).$$

The identity (50) is thus proved. \square

Lemma 4.5. *For any $\sigma \geq \tau$ and for all $f \in \mathcal{K}^\infty(\pi_\Lambda^X)$,*

$$(52) \quad |G_\Lambda^X(f)|_{\mathcal{F}, \tau} \leq J_\sigma^\tau(\Lambda, \mathcal{F}) |f|_{\mathcal{F}, \sigma}.$$

Proof. It follows by Hölder inequality from formula (48) for the Green operator, in fact, for all $x \in \mathbb{R}$, by Hölder inequality we have

$$|G_\Lambda^X(f)(x)|^2 \leq \left(\int_{|y| \geq |x|} \frac{dy}{(1 + \Delta_{\Lambda, \mathcal{F}}(y))^\sigma} \right) |f|_{\mathcal{F}, \sigma}^2.$$

Another application of Hölder inequality yields the result. \square

We have thus reduced Sobolev bounds on the Green operator for the cohomological equation and on the ergodic averages operator (in each irreducible representation) to bounds on the integrals defined in formula (49).

Let $\mathcal{F} = (X, Y_1, \dots, Y_a)$ be any adapted basis. Let $(d_1, \dots, d_a) \in \mathbb{N}^a$ be the degrees of the elements (Y_1, \dots, Y_a) , respectively; for all $\Lambda \in \mathfrak{a}_0^*$, let $\Lambda_i^{(j)}(\mathcal{F}) = \Lambda(\text{ad}(X)^j Y_i)$ be the coefficients appearing in formula (38) and set

$$(53) \quad |\Lambda(\mathcal{F})| := \sup_{\{(i, j) : 1 \leq i \leq a, 0 \leq j \leq d_i\}} \left| \frac{1}{j!} \Lambda_i^{(j)}(\mathcal{F}) \right|.$$

We introduce on \mathfrak{a}_0^* the following weight. For all $\Lambda \in \mathfrak{a}_0^*$, let

$$(54) \quad w_{\mathcal{F}}(\Lambda) := \min_{\{i : d_i \neq 0\}} \left| \frac{\Lambda_i^{(d_i)}(\mathcal{F})}{d_i!} \right|^{-\frac{1}{d_i}}$$

We will prove below estimates for the integrals $I_\sigma(\Lambda, \mathcal{F})$ and $J_\sigma^\tau(\Lambda, \mathcal{F})$ of formula (49) in terms of the above weight.

For all $i = 1, \dots, a$ and $j = 1, \dots, d_i$ we define the rescaled coefficients

$$(55) \quad \hat{\Lambda}_i^{(j)}(\mathcal{F}) := \Lambda_i^{(j)}(\mathcal{F}) (w_{\mathcal{F}}(\Lambda))^j,$$

and set, in analogy with (53),

$$(56) \quad |\hat{\Lambda}(\mathcal{F})| := \sup_{\{(i, j) : 1 \leq i \leq a, 0 \leq j \leq d_i\}} \left| \frac{1}{j!} \hat{\Lambda}_i^{(j)}(\mathcal{F}) \right|.$$

Lemma 4.6. *For all $\sigma > 1/2$, there exists a constant $C_{k,\sigma} > 0$ such that, for all $\Lambda \in \mathfrak{a}_0^*$, the following bounds hold:*

$$(57) \quad \frac{C_{k,\sigma}^{-1}}{(1 + |\hat{\Lambda}(\mathcal{F})|)^\sigma} \leq \frac{I_\sigma(\Lambda, \mathcal{F})}{w_{\mathcal{F}}^{1/2}(\Lambda)} \leq C_{k,\sigma}(1 + |\hat{\Lambda}(\mathcal{F})|).$$

For all $\sigma > \tau(k-1) + 1$, there exists a constant $C_{k,\sigma,\tau} > 0$ such that, for all $\Lambda \in \mathfrak{a}_0^$, the following bounds hold:*

$$(58) \quad \frac{C_{k,\sigma,\tau}^{-1}}{(1 + |\hat{\Lambda}(\mathcal{F})|)^\sigma} \leq \frac{J_\sigma^\tau(\Lambda, \mathcal{F})}{w_{\mathcal{F}}(\Lambda)} \leq C_{k,\sigma,\tau}(1 + |\hat{\Lambda}(\mathcal{F})|)^{\tau k + 1}.$$

Proof. By change of variables, for any $w > 0$,

$$(59) \quad \begin{aligned} I_\sigma(\Lambda, \mathcal{F}) &= w^{1/2} \left(\int_{\mathbb{R}} \frac{dx}{[1 + \Delta_{\Lambda, \mathcal{F}}(wx)]^\sigma} \right)^{1/2}; \\ J_\sigma^\tau(\Lambda, \mathcal{F}) &= w \left(\int \int_{|y| \geq |x|} \frac{[1 + \Delta_{\Lambda, \mathcal{F}}(wx)]^\tau}{[1 + \Delta_{\Lambda, \mathcal{F}}(wy)]^\sigma} dx dy \right)^{1/2}. \end{aligned}$$

Let $w := w_{\mathcal{F}}(\Lambda) > 0$. By definitions (54), (55) and (44), for all $i = 1, \dots, a$, the coefficients of the polynomial map $P_{\Lambda, Y_i}(wx)$ are the numbers $\hat{\Lambda}_i^{(j)}(\mathcal{F})/j!$. Thus all these coefficients are bounded by $|\hat{\Lambda}(\mathcal{F})|$ and there exists $i_0 \in \{1, \dots, a\}$ such that the polynomial $P_{\Lambda, Y_{i_0}}(wx)$ is monic. The following inequalities therefore hold: for all $x \in \mathbb{R}$,

$$(60) \quad 1 + P_{\Lambda, Y_{i_0}}^2(wx) \leq 1 + \Delta_{\Lambda, \mathcal{F}}(wx) \leq (1 + |\hat{\Lambda}(\mathcal{F})|)^2(1 + x^{2(k-1)}).$$

Let $P(x)$ be any non-constant monic polynomial of degree $d \geq 1$ and let $\|P\|$ denote the maximum modulus of its coefficients. We claim that, if $d\sigma > 1/2$, there exists a constant $C_{d,\sigma} > 0$ such that

$$(61) \quad \int_{\mathbb{R}} \frac{dx}{(1 + P^2(x))^\sigma} \leq C_{d,\sigma}(1 + \|P\|),$$

and, if $d\sigma > (k-1)\tau + 1/2$, there exists a constant $C_{k,d,\sigma,\tau} > 0$ such that

$$(62) \quad \int \int_{|y| \geq |x|} \frac{(1 + x^{2(k-1)})^\tau}{(1 + P^2(y))^\sigma} dx dy \leq C_{k,d,\sigma,\tau}(1 + \|P\|)^{2+2\tau(k-1)}.$$

In fact, since P is monic, there exists $s \in [1, (1 + \|P\|)]$ such that the polynomial $P_s(x) := s^{-d}P(sx)$ is monic and has all coefficients in the unit ball. It follows that, if $d\sigma > 1/2$, there exists a constant $C_{d,\sigma} > 0$ such that

$$\int_{\mathbb{R}} \frac{dx}{(1 + P^2(x))^\sigma} = \int_{\mathbb{R}} \frac{s dx}{(1 + P^2(sx))^\sigma} \leq \int_{\mathbb{R}} \frac{s dx}{(1 + P_s^2(x))^\sigma} \leq C_{d,\sigma}s,$$

hence the bound in formula (61) is proved. Similarly, if $\sigma > \tau(k-1) + 1$, there exists a constant $C_{k,d,\sigma,\tau} > 0$ such that

$$\begin{aligned} \int \int_{|y| \geq |x|} \frac{(1 + x^{2(k-1)})^\tau}{(1 + P^2(y))^\sigma} dx dy &= \int \int_{|y| \geq |x|} s^2 \frac{(1 + (sx)^{2(k-1)})^\tau}{(1 + P^2(sy))^\sigma} dx dy \\ &\leq \int \int_{|y| \geq |x|} s^2 \frac{(1 + (sx)^{2(k-1)})^\tau}{(1 + P_s^2(sy))^\sigma} dx dy \leq C_{k,d,\sigma,\tau} s^{2+2\tau(k-1)}, \end{aligned}$$

hence the bound in formula (62) is proved as well.

Finally, applying the the bounds in (61) and (62) to the polynomial $P_{\Lambda, Y_i}(wx)$ and taking into account the formulas (59) and the estimates (60) we obtain the upper bounds (57) and (58).

The lower bounds are an immediate consequence of the upper bound in formula (60), hence the argument is complete. \square

4.3. The renormalisation group.

Definition 4.7. The *deformation space* of a k -step nilpotent quasi-Abelian nilmanifold $M = \Gamma \backslash G$ is the space $T(M)$ of all adapted bases of the Lie algebra \mathfrak{g} of the group G .

Let $\mathcal{A} < SL(a+1, \mathbb{R})$ be the subgroup of all matrices A of the following form. For any $\alpha \in \mathbb{R} \setminus \{0\}$, for any vector $\beta \in \mathbb{R}^a$ and for any matrix $B \in GL(a, \mathbb{R})$, let

$$(63) \quad A := \begin{pmatrix} \alpha & \beta \\ 0 & B \end{pmatrix}.$$

The group \mathcal{A} acts on the deformation space $T(M)$. In fact, let $\mathcal{F} = (X, Y)$ be any adapted basis of the Lie algebra \mathfrak{g} . For any $A \in \mathcal{A}$ the transformed basis is defined as

$$(64) \quad \begin{pmatrix} X_A \\ Y_{1,A} \\ \dots \\ Y_{a,A} \end{pmatrix} = A \begin{pmatrix} X \\ Y_1 \\ \dots \\ Y_a \end{pmatrix}.$$

The renormalisation dynamics will be defined as the action of the diagonal subgroup of the Lie group $SL(a+1, \mathbb{R})$ on the deformation space.

Let $\rho := (\rho_1, \dots, \rho_a) \in (\mathbb{R}^+)^a$ be any vector such that

$$\sum_{j=1}^a \rho_j = 1,$$

there exists a one-parameter subgroup $\{A_t^\rho\}$ of the diagonal subgroup of $SL(a+1, \mathbb{R})$ defined as follows:

$$(65) \quad A_t^\rho(X, \dots, Y_i, \dots) = (e^t X, \dots, e^{-\rho_i t} Y_i, \dots).$$

Note that the renormalisation group preserves the set of all generalised Jordan basis. However, the group \mathcal{A} is not a group of automorphisms of the Lie algebra. Consequently, the dynamics induced by the renormalisation group on the deformation space is trivial (it has no recurrent orbits).

4.3.1. Estimates for rescaled bases. In this section we prove Sobolev estimates for the invariant distribution and for the Green operator in any irreducible unitary representation with respect to rescaled bases. Let $\Lambda \in \mathfrak{a}_0^*$, let $\mathcal{F} = (X, Y)$ be any adapted basis and let π_Λ^X be the induced representation. For all $t \in \mathbb{R}$, let

$$(66) \quad \mathcal{F}(t) = (X(t), Y(t)) = A_t^\rho(X, Y)$$

a rescaled adapted basis and let $U_t : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ be the unitary operator defined as follows: for all $f \in L^2(\mathbb{R})$,

$$(U_t f)(x) = e^{-\frac{t}{2}} f(e^{-t} x), \quad \text{for all } x \in \mathbb{R}.$$

Lemma 4.8. *For all $t \in \mathbb{R}$ the operator $U_t : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ intertwines the representation $\pi_\Lambda^{X(t)}$ and π_Λ^X :*

$$U_t^{-1} \pi_\Lambda^X U_t = \pi_\Lambda^{X(t)}.$$

Proof. By simple computations it follows that, for all $f \in L^2(\mathbb{R})$,

$$[U_t^{-1} \pi_\Lambda^X(X(t)) U_t](f) = \frac{df}{dx} = \pi_\Lambda^{X(t)}(X(t))(f),$$

and, for all $i = 1, \dots, a$ and for all $x \in \mathbb{R}$,

$$\begin{aligned} U_t^{-1} \pi_\Lambda^X(Y_i) U_t f(x) &= i \left[\sum_{j \geq 0} \frac{(\Lambda \circ \text{ad}_{X(t)}^j)(Y_i)}{j!} e^{jt} x^j \right] f(x) \\ &= i \left[\sum_{j \geq 0} \frac{(\Lambda \circ \text{ad}_{X(t)}^j)(Y_i)}{j!} x^j \right] f(x) = \pi_\Lambda^{X(t)}(Y_i)(f)(x). \end{aligned}$$

It follows that the representations $U_t^{-1} \pi_\Lambda^X U_t$ and $\pi_\Lambda^{X(t)}$ are equal since they coincide on the basis $(X(t), Y)$ of the Lie algebra. \square

The Sobolev norms of the invariant distribution and of the Green operator with respect to the rescaled basis $\mathcal{F}(t)$ are given below.

Lemma 4.9. *For $\sigma > 1/2$ and for all $t \in \mathbb{R}$, the following holds:*

$$|\mathcal{D}_\Lambda^X|_{\mathcal{F}(t), -\sigma} = e^{\frac{t}{2}} I_\sigma(\Lambda, \mathcal{F}(t)).$$

Proof. By Lemma 4.8, by change of variable we have that, for all $f \in L^2(\mathbb{R})$,

$$\mathcal{D}_\Lambda^X(f) = \mathcal{D}_\Lambda^{X(t)}(U_t f) = e^{\frac{t}{2}} \mathcal{D}_\Lambda^{X(t)}(f).$$

It follows that, by Lemma 4.4, we have

$$|\mathcal{D}_\Lambda^X|_{\mathcal{F}(t), -\sigma} = e^{\frac{t}{2}} |\mathcal{D}_\Lambda^{X(t)}|_{\mathcal{F}(t), -\sigma} = e^{\frac{t}{2}} I_\sigma(\Lambda, \mathcal{F}(t)).$$

\square

Let $G_{X,\Lambda}^{X(t)}$ denote the Green operator for the cohomological equation $X(t)u = f$ in the representation π_Λ^X . We recall that, according to our definitions above, see formula (48), $G_\Lambda^{X(t)}$ denote the Green operator for the same cohomological equation $X(t)u = f$ in the representation $\pi_\Lambda^{X(t)}$.

Lemma 4.10. *For $\sigma > \tau$ and for all $t \in \mathbb{R}$, for all $f \in \mathcal{K}^\infty(\pi_\Lambda^{X(t)})$, the following holds:*

$$|G_{X,\Lambda}^{X(t)}(f)|_{\mathcal{F}(t), \tau} \leq J_\sigma^\tau(\Lambda, \mathcal{F}(t)) |f|_{\mathcal{F}(t), \sigma}.$$

Proof. By Lemma 4.8 the operators $G_{X,\Lambda}^{X(t)}$ and $G_\Lambda^{X(t)}$ are unitarily equivalent:

$$G_{X,\Lambda}^{X(t)} = U_t \circ G_\Lambda^{X(t)} \circ U_t^{-1}$$

hence by Lemma 4.5 we have, for all $f \in L^2(\mathbb{R})$,

$$\begin{aligned} |G_{X,\Lambda}^{X(t)}(f)|_{\mathcal{F}(t), \tau} &= |U_t^{-1} G_\Lambda^{X(t)}(U_t f)|_{\mathcal{F}(t), \tau} = |G_\Lambda^{X(t)}(U_t f)|_{\mathcal{F}(t), \tau} \\ &\leq J_\sigma^\tau(\Lambda, \mathcal{F}(t)) |U_t f|_{\mathcal{F}(t), \sigma} = J_\sigma^\tau(\Lambda, \mathcal{F}(t)) |f|_{\mathcal{F}(t), \sigma}. \end{aligned}$$

\square

Sobolev estimates of the normalised invariant distributions and of Green operators for rescaled bases are thus reduced to bounds on the integrals $I_\sigma(\Lambda, \mathcal{F}(t))$ and $J_\sigma^\tau(\Lambda, \mathcal{F}(t))$, which by Lemma 4.6 can be bounded uniformly in terms of the weights $w_{\mathcal{F}(t)}(\Lambda)$. We therefore proceed to estimate the latter.

Let $\mathcal{F} = (X, Y_1, \dots, Y_a)$ be any adapted basis and let $(d_1, \dots, d_a) \in \mathbb{N}^a$ denote the vector of the degrees of the elements (Y_1, \dots, Y_a) respectively. For any $\rho = (\rho_1, \dots, \rho_a) \in (\mathbb{R}^+)^a$, let

$$(67) \quad \lambda_{\mathcal{F}}(\rho) := \min_{\{i: d_i \neq 0\}} \left(\frac{\rho_i}{d_i} \right)$$

Lemma 4.11. *For any adapted basis $\mathcal{F} = (X, Y)$ and for all $t \geq 0$,*

$$(68) \quad e^{-(1-\lambda_{\mathcal{F}}(\rho))t} w_{\mathcal{F}}(\Lambda) \leq w_{\mathcal{F}(t)}(\Lambda) \leq \left(\max_{\{i: d_i \neq 0\}} |\Lambda_i^{(d_i)}(\mathcal{F})|^{-\frac{1}{d_i}} \right) e^{-(1-\lambda_{\mathcal{F}}(\rho))t}.$$

Proof. Since

$$|\Lambda(\text{ad}_{X(t)}^{d_i} Y_i(t))|^{-\frac{1}{d_i}} = e^{-t(1-\rho_i/d_i)} |\Lambda(\text{ad}_X^{d_i} Y_i)|^{-\frac{1}{d_i}}$$

the two inequalities follow immediately from the definition (54) and the above definition of $\lambda_{\mathcal{F}}(\rho)$. \square

We also estimate the normalised coefficients $\hat{\Lambda}(\mathcal{F}(t))$ of the representation. For convenience of notation we introduce the following weight: for all $\Lambda \in \mathfrak{a}_0^*$, let

$$(69) \quad \|\Lambda\|_{\mathcal{F}} := |\Lambda(\mathcal{F})| \max_{\{i: d_i \neq 0\}} \left(1 + \frac{1}{|\Lambda_i^{(d_i)}(\mathcal{F})|} \right).$$

Lemma 4.12. *For any adapted basis $\mathcal{F} = (X, Y)$ and for all $i \in \{1, \dots, a\}$ the following bound holds: for all $t \geq 0$,*

$$(70) \quad |\hat{\Lambda}(\mathcal{F}(t))| \leq \|\Lambda\|_{\mathcal{F}}.$$

Proof. By the definitions (37), (55), (65) and (66), we have, for all $i = 1, \dots, a$ and $j = 1, \dots, d_i$

$$\hat{\Lambda}_i^{(j)}(\mathcal{F}(t)) = \Lambda_i^{(j)}(\mathcal{F}(t)) (w_{\mathcal{F}(t)}(\Lambda))^j = e^{(j-\rho_i)t} \Lambda_i^{(j)}(\mathcal{F}) (w_{\mathcal{F}(t)}(\Lambda))^j.$$

By Lemma 4.11 and observing that $j \leq d_i$ we obtain

$$|\hat{\Lambda}_i^{(j)}(\mathcal{F}(t))| \leq |\Lambda_i^{(j)}(\mathcal{F})| \max_{\{i: d_i \neq 0\}} |\Lambda_i^{(d_i)}(\mathcal{F})|^{-\frac{j}{d_i}};$$

the bound (70) follows from the elementary estimate

$$|\Lambda_i^{(d_i)}(\mathcal{F})|^{-\frac{j}{d_i}} \leq \left(1 + \frac{1}{|\Lambda_i^{(d_i)}(\mathcal{F})|} \right),$$

and from the definitions (53) and (56). \square

We finally conclude the section with the fundamental estimates on the scaling of invariant distributions of the Green operator.

Theorem 4.13. *For all $\sigma > 1/2$, there exists a constant $D_{k,\sigma} > 0$ such that, for all $t \geq 0$,*

$$|\mathcal{D}_\Lambda^X|_{\mathcal{F}, -\sigma} \leq D_{k,\sigma} (1 + \|\Lambda\|_{\mathcal{F}})^{\sigma+1} e^{-\frac{\lambda_{\mathcal{F}}(\rho)}{2}t} |\mathcal{D}_\Lambda^X|_{\mathcal{F}(t), -\sigma}.$$

Proof. By Lemma 4.4 and Lemma 4.9 we have that

$$(71) \quad \frac{|\mathcal{D}_\Lambda^X|_{\mathcal{F}, -\sigma}}{|\mathcal{D}_\Lambda^X|_{\mathcal{F}(t), -\sigma}} = e^{-\frac{t}{2}} \frac{I_\sigma(\Lambda, \mathcal{F})}{I_\sigma(\Lambda, \mathcal{F}(t))}.$$

By Lemma 4.6, Lemma 4.11 and Lemma 4.12 we obtain the estimate

$$(72) \quad \begin{aligned} \frac{I_\sigma(\Lambda, \mathcal{F})}{I_\sigma(\Lambda, \mathcal{F}(t))} &\leq C_{k, \sigma}^2 (1 + |\hat{\Lambda}(\mathcal{F})|)(1 + |\hat{\Lambda}(\mathcal{F}(t))|)^\sigma \left(\frac{w_{\mathcal{F}}(\Lambda)}{w_{\mathcal{F}(t)}(\Lambda)} \right)^{1/2} \\ &\leq C_{k, \sigma}^2 (1 + |\Lambda(\mathcal{F})|)^{\sigma+1} \max_{\{i: d_i \neq 0\}} \left(1 + \frac{1}{\Lambda_i^{(d_i)}(\mathcal{F})} \right)^{\sigma+1} e^{\frac{1-\lambda_{\mathcal{F}}(\rho)}{2}t}. \end{aligned}$$

And the statement follows. \square

Theorem 4.14. *For $\sigma > \tau(k-1) + 1$ there exists a constant $G_{k, \sigma, \tau} > 0$ such that, for all $t \in \mathbb{R}$ and for all $f \in \mathcal{K}^\infty(\pi_\Lambda^{X(t)})$, the following holds:*

$$(73) \quad |G_{X, \Lambda}^{X(t)}(f)|_{\mathcal{F}(t), \tau} \leq G_{k, \sigma, \tau} (1 + \|\Lambda\|_{\mathcal{F}})^{\tau k + 2} e^{-(1-\lambda_{\mathcal{F}}(\rho))t} |f|_{\mathcal{F}(t), \sigma}.$$

Proof. It is an immediate consequence of the estimate (58) of Lemma 4.6 and of Lemma 4.11. \square

4.3.2. A Lyapunov norm. For convenience we introduce a Lyapunov norm on the space of invariant distributions in each irreducible unitary representation. For any adapted basis \mathcal{F} , for all $\Lambda \in \mathfrak{a}_0^*$ and for all $\sigma > 1/2$, let

$$(74) \quad \|\mathcal{D}_\Lambda^X\|_{\mathcal{F}, -\sigma} := \inf_{\tau \geq 0} e^{-\frac{\lambda_{\mathcal{F}}(\rho)}{2}\tau} |\mathcal{D}_\Lambda^X|_{\mathcal{F}(\tau), -\sigma}$$

It follows from the definition and from Theorem 4.13 that

$$(75) \quad \frac{|\mathcal{D}_\Lambda^X|_{\mathcal{F}, -\sigma}}{D_{k, \sigma}(1 + \|\Lambda\|_{\mathcal{F}})^{\sigma+1}} \leq \|\mathcal{D}_\Lambda^X\|_{\mathcal{F}(t), -\sigma} \leq |\mathcal{D}_\Lambda^X|_{\mathcal{F}, -\sigma}.$$

Lemma 4.15. *For all $t \geq 0$, we have*

$$\|\mathcal{D}_\Lambda^X\|_{\mathcal{F}, -\sigma} \leq e^{-\frac{\lambda_{\mathcal{F}}(\rho)}{2}t} \|\mathcal{D}_\Lambda^X\|_{\mathcal{F}(t), -\sigma}.$$

Proof. It follows immediately from the definition of the norm. In fact,

$$\begin{aligned} \|\mathcal{D}_\Lambda^X\|_{\mathcal{F}, -\sigma} &= \inf_{\tau \geq 0} e^{-\frac{\lambda_{\mathcal{F}}(\rho)}{2}\tau} |\mathcal{D}_\Lambda^X|_{\mathcal{F}(\tau), -\sigma} \\ &= e^{-\frac{\lambda_{\mathcal{F}}(\rho)}{2}t} \inf_{t+\tau \geq 0} e^{-\frac{\lambda_{\mathcal{F}}(\rho)}{2}\tau} |\mathcal{D}_\Lambda^X|_{\mathcal{F}(t+\tau), -\sigma} \leq e^{-\frac{\lambda_{\mathcal{F}}(\rho)}{2}t} \|\mathcal{D}_\Lambda^X\|_{\mathcal{F}(t), -\sigma}. \end{aligned}$$

\square

5. AVERAGE WIDTH ESTIMATES

In this section we prove estimates on the average width of orbits of quasi-Abelian nilflows. Let $\alpha := (\alpha_i^{(m)}) \in \mathbb{R}^J$ and let X_α be the vector field on M defined in formula (16), that is

$$X_\alpha := \log \left[x^{-1} \exp \left(\sum_{(m, i) \in J} \alpha_i^{(m)} \tilde{\eta}_i^{(m)} \right) \right].$$

Let $\{\phi_{X_\alpha}^t\}$ be the flow generated by the vector field X_α on M .

5.1. Almost periodic points. Let us recall that \mathbb{T}_θ^a denotes the fibre at $\theta \in \mathbb{T}^1$ of the fibration $\text{pr}_2 : M \rightarrow \mathbb{T}^1$ (see formula (13)), that $\Phi_{\alpha,\theta}$ denotes the first return map of the quasi-Abelian nilflow $\{\phi_{X_\alpha}^t\}$ to the transverse torus \mathbb{T}_θ^a (see Lemma 2.4) and that $\Phi_{\alpha,\theta}^r : \mathbb{T}_\theta^a \rightarrow \mathbb{T}_\theta^a$ denotes, for all $r \in \mathbb{Z}$, the r -th iterate of the map $\Phi_{\alpha,\theta}$. By its definition, the map $\Phi_{\alpha,\theta}$ commutes with the action of the centre $Z(G)$ of quasi-Abelian nilpotent group G ; hence, for all $r \in \mathbb{Z}$, the map $\Phi_{\alpha,\theta}^r - \text{Id}$ induces a quotient map

$$\Psi_{\alpha,\theta}^{(r)} : \mathbb{T}_\theta^a / Z(G) \rightarrow \mathbb{T}_\theta^a.$$

For every $m = 1, \dots, n$ and for all $\theta \in \mathbb{T}^1$, let $\mathbb{Z}_\theta^{i_m} \subset \mathbb{R}^{i_m}$ be the lattice introduced in formula (17) and let $\mathbb{T}_\theta^{i_m} \subset \mathbb{T}^a$ be the related sub-torus introduced in formula (18). By Lemma 2.4 the map $\Psi_{\alpha,\theta}^{(r)}$ has a factorisation

$$\Psi_{\alpha,\theta}^{(r)} \approx \Psi_{\alpha^{(1)},\theta}^{(r)} \times \dots \times \Psi_{\alpha^{(n)},\theta}^{(r)} \quad \text{on} \quad (\mathbb{T}_\theta^{i_1} \times \dots \times \mathbb{T}_\theta^{i_n}) / Z(G),$$

and, for every $m = 1, \dots, n$, the factor map $\Psi_{\alpha^{(m)},\theta}^{(r)}$ is given in the coordinates $\mathbf{s}^{(m)} \in \mathbb{R}^{i_m} \bmod \mathbb{Z}_\theta^{i_m}$ by the formulas

$$(76) \quad \begin{aligned} \Psi_{\alpha^{(m)},\theta}^{(r)}(\mathbf{s}^{(m)}) &= (r \alpha_1^{(m)}, r(s_1^{(m)} + \alpha_2^{(m)}) + \binom{r}{2} \alpha_1^{(m)}, \dots, \\ &\quad \sum_{i=1}^{i_m-1} \binom{r}{i} (s_{i_m-i}^{(m)} + \alpha_{i_m-i+1}^{(m)}) + \binom{r}{i_m} \alpha_1^{(m)}). \end{aligned}$$

It is clear that the above formulas define a map on the quotient $\mathbb{T}_\theta^a / Z(G)$. In fact, for all $m = 1, \dots, n$, the map $\Psi_{\alpha^{(m)},\theta}^{(r)}$ does not depend on the coordinate $s_{i_m}^{(m)} \in \mathbb{R}$. It is also clear that the first coordinate of the map $\Psi_{\alpha^{(m)},\theta}^{(r)}$ is constant (equal to $r \alpha_1^{(m)}$), hence the image of the map $\Psi_{\alpha,\theta}^{(r)}$ is contained in the affine $(a-n)$ -dimensional sub-torus

$$\mathbb{T}_{\theta,\alpha,r}^{a-n} := \text{pr}_1^{-1} \{ \text{pr}_1(\Gamma \exp(\theta \xi) \exp(r \sum_{m=1}^n \alpha_1^{(m)} \tilde{\eta}_1^{(m)})) \} \subset \mathbb{T}_\theta^a.$$

We recall that, according to formulas (12) the map $\text{pr}_1 : M \rightarrow M_1 \approx \mathbb{T}^{n+1}$ is the projection on the base torus. Let \mathcal{L}_θ^{a-n} denote the $(a-n)$ -dimensional conditional Lebesgue measure on the torus $\mathbb{T}_\theta^a / Z(G)$. By construction the coordinates

$$(s_1^{(1)}, \dots, s_{i_1-1}^{(1)}, \dots, s_1^{(n)}, \dots, s_{i_n-1}^{(n)}) \in \mathbb{R}^{a-n},$$

taken modulo the action of the lattice $\mathbb{Z}_\theta^a \cap \{s_{i_1}^{(1)} = \dots = s_{i_n}^{(n)} = 0\}$, are well-defined coordinates for the quotient torus $\mathbb{T}_\theta^a / Z(G)$; by the above discussion, the measure \mathcal{L}_θ^{a-n} can be written in coordinates as follows:

$$\mathcal{L}_\theta^{a-n} = ds_1^{(1)} \dots ds_{i_1-1}^{(1)} \dots ds_1^{(n)} \dots ds_{i_n-1}^{(n)}.$$

Similarly, let $\mathcal{L}_{\theta,\alpha,r}^{a-n}$ denote $(a-n)$ -dimensional conditional Lebesgue measure on the torus $\mathbb{T}_{\theta,\alpha,r}^{a-n}$. The coordinates

$$(s_2^{(1)}, \dots, s_{i_1}^{(1)}, \dots, s_2^{(n)}, \dots, s_{i_1}^{(n)}) \in \mathbb{R}^{a-n},$$

taken modulo the action of the lattice $\mathbb{Z}_\theta^a \cap \{s_1^{(1)} = \dots = s_1^{(n)} = 0\}$ are well-defined coordinates for the sub-torus $\mathbb{T}_{\theta,\alpha,r}^{a-n}$; the measure $\mathcal{L}_{\theta,\alpha,r}^{a-n}$ can be written in

coordinates as follows:

$$\mathcal{L}_{\theta, \alpha, r}^{a-n} = ds_2^{(1)} \dots ds_{i_1}^{(1)} \dots ds_2^{(n)} \dots ds_{i_n}^{(n)}.$$

Lemma 5.1. *For all $\alpha \in \mathbb{R}^J$ and all $\theta \in \mathbb{T}^1$ the map $\Psi_{\alpha, \theta}^{(r)}$ is a covering map of the torus $\mathbb{T}_\theta^a/Z(G)$ onto the torus $\mathbb{T}_{\theta, \alpha, r}^{a-n}$ with constant Jacobian. Hence it maps the measure \mathcal{L}_θ^{a-n} onto the measure $\mathcal{L}_{\theta, \alpha, r}^{a-n}$.*

Proof. By formula (76), for any $\alpha \in \mathbb{R}^J$, for every $m = 1, \dots, n$ and for every $j \in \{2, \dots, i_m\}$, there exists a polynomial $p_j^{(m)}(\alpha, r, s_1^{(m)}, \dots, s_{j-2}^{(m)})$ such that the j -th coordinate $\Psi_{\alpha^{(m)}, \theta, j}^{(r)}$ of the map $\Psi_{\alpha^{(m)}, \theta}^{(r)}$ is given by the following formula: for all $\mathbf{s}^{(m)} \in \mathbb{R}^{i_m}$,

$$\Psi_{\alpha^{(m)}, \theta, j}^{(r)}(\mathbf{s}^{(m)}) = rs_{j-1}^{(m)} + p_j(\alpha, r, s_1^{(m)}, \dots, s_{j-2}^{(m)}).$$

It follows that the Jacobian of $\Psi_{\alpha, \theta}^{(r)}$ is a non zero constant and the map is a regular covering. Hence the push-forward of the Lebesgue measure \mathcal{L}_θ^{a-n} on $\mathbb{T}_\theta^a/Z(G)$ to the torus $\mathbb{T}_{\theta, \alpha, r}^{a-n}$ under $\Psi_{\alpha, \theta}^{(r)}$ is the Lebesgue measure $\mathcal{L}_{\theta, \alpha, r}^{a-n}$. \square

Let \mathcal{U}_θ be a given neighbourhood of the origin in \mathbb{R}^a . A point $x \in \mathbb{T}_\theta^a$ is (\mathcal{U}_θ, r) -almost-periodic, that is, it is \mathcal{U}_θ -almost-periodic of period $r \in \mathbb{N}$ for $\Phi_{\alpha, \theta}$ if $\Phi_{\alpha, \theta}^r(x)$ belongs to the neighbourhood $x + \mathcal{U}_\theta$ of $x \in \mathbb{T}_\theta^a$. For every $r \in \mathbb{Z} \setminus \{0\}$, let $\text{AP}^r(\mathcal{U}_\theta)$ be the set of (\mathcal{U}_θ, r) -almost-periodic points:

$$\text{AP}^r(\mathcal{U}_\theta) := \{x \in \mathbb{T}_\theta^a \mid \Phi_{\alpha, \theta}^r(x) - x \in \mathcal{U}_\theta\}.$$

The next lemma estimates the Lebesgue measure of the set $\text{AP}^r(\mathcal{U}_\theta)$ of (\mathcal{U}_θ, r) -almost-periodic points. To this purpose let us introduce yet another projection map: for all $\theta \in \mathbb{T}^1$, let $\text{pr}_\theta : \mathbb{T}_\theta^a \rightarrow \mathbb{T}^n$ be the restriction to the torus \mathbb{T}_θ^a of the projection $\text{pr}_1 : M \rightarrow \mathbb{T}^{n+1}$ onto the base torus, that is, the map defined by the following formula: for all $\mathbf{s} \in \mathbb{R}^a$,

$$\text{pr}_\theta : \Gamma \exp(\theta \xi) \exp \left(\sum_{m=1}^n \sum_{i=1}^{i_m} s_{i_m}^{(m)} \tilde{\eta}_{i_m}^{(m)} \right) = (s_1^{(1)}, \dots, s^{(n)}) \pmod{\mathbb{Z}^n}.$$

Lemma 5.2. *Let $\theta \in \mathbb{T}^1$ and let $\mathcal{U}_\theta \subset \mathbb{T}_\theta^a$ be any neighbourhood of the origin. For all $r \in \mathbb{Z} \setminus \{0\}$, the a -dimensional conditional Lebesgue measure \mathcal{L}_θ^a of the set $\text{AP}^r(\mathcal{U}_\theta) \subset \mathbb{T}_\theta^a$ is given as follows.*

If $r\alpha_1 \in \text{pr}_\theta(\mathcal{U}_\theta)$, then

$$\mathcal{L}_\theta^a(\text{AP}^r(\mathcal{U}_\theta)) = \mathcal{L}_{\theta, \alpha, r}^{a-n}(\mathcal{U}_\theta \cap \mathbb{T}_{\theta, \alpha, r}^{a-n});$$

otherwise $\text{AP}^r(\mathcal{U}_\theta) = \emptyset$.

Proof. The sets $\text{AP}^r(\mathcal{U}_\theta)$ are invariant under the action of the centre $Z(G)$. By definition, the projection of a set $\text{AP}^r(\mathcal{U}_\theta)/Z(G)$ to the quotient torus $\mathbb{T}_\theta^a/Z(G)$ is the inverse image of the neighbourhood $\mathcal{U}_\theta \subset \mathbb{T}_\theta^a$ under the map defined on $\mathbb{T}_\theta^a/Z(G)$ as

$$x \pmod{Z(G)} \mapsto (r\alpha_1, \Psi_{\alpha, \theta}^{(r)}(x)) \in \mathbb{T}_\theta^a.$$

Thus $\text{AP}^r(\mathcal{U}_\theta) = \emptyset$ if $r\alpha_1 \notin \text{pr}_\theta(\mathcal{U}_\theta)$; if $r\alpha_1 \in \text{pr}_\theta(\mathcal{U}_\theta)$, then

$$\text{AP}^r(\mathcal{U}_\theta)/Z(G) = (\Psi_{\alpha, \theta}^{(r)})^{-1}(\mathcal{U}_\theta \cap \mathbb{T}_{\theta, \alpha, r}^{a-n}).$$

The result then follows from Lemma 5.1. \square

Definition 5.3. For any basis $Y = \{Y_1, \dots, Y_a\}$ of the Abelian ideal $\mathfrak{a} \subset \mathfrak{g}$, let $I := I(Y)$ be the supremum of all constants $I' \in]0, 1/2[$ such that for any $x \in M$ the map

$$\phi_x^Y : (s_1, \dots, s_a) \rightarrow x \exp\left(\sum_{i=1}^a s_i Y_i\right) \in M.$$

is a local embedding (injective) on the domain

$$\{\mathbf{s} \in \mathbb{R}^a \mid |s_i| < I' \text{ for all } i = 1, \dots, a\}.$$

For any $x, x' \in M$, we set

$$\|x' - x\|_1 = |s_1|, \dots, \|x' - x\|_i = |s_i|, \dots, \|x' - x\|_a = |s_a|,$$

if there is $\mathbf{s} := (s_1, \dots, s_a) \in [-I/2, I/2]^a$ such that

$$x' = x \exp\left(\sum_{i=1}^a s_i Y_i\right);$$

otherwise we set $\|x' - x\|_1 = \dots = \|x' - x\|_a = I$.

Let $\rho := (\rho_1, \dots, \rho_a) \in [0, 1]^a$ and let $\mathcal{F}_\alpha = (X_\alpha, Y)$ be a normalised *strongly adapted basis* (see Definition 3.1). Let us observe that, since the basis \mathcal{F}_α is strongly adapted, the vector

$$(\|\Phi_{\alpha, \theta}^r(x) - x\|_1, \dots, \|\Phi_{\alpha, \theta}^r(x) - x\|_n)$$

does not depend upon the choice of $x \in M$, but only depends on $r \in \mathbb{Z}$; in fact, the subsystem (Y_{n+1}, \dots, Y_a) is tangent to the fibres of the projection $\text{pr}_\theta : \mathbb{T}_\theta^a \rightarrow \mathbb{T}^n$, and, for all $x \in \mathbb{T}_\theta^a$, we have

$$\text{pr}_\theta(\Phi_{\alpha, \theta}^r(x) - x) = r\alpha_1 \mod \mathbb{Z}^n.$$

It follows that, for any $L \geq 1$ and for any $r \in \mathbb{Z}$, we can define

$$(77) \quad \epsilon_{r, L} := \max_{1 \leq i \leq n} \min\{I, L^{\rho_i} \|\Phi_{\alpha, \theta}^r(x) - x\|_i\}.$$

For $L \geq 1$, $r \in \mathbb{Z}$ and $x \in \mathbb{T}_\theta^a$, we also define

$$\delta_{r, L}(x) := \max_{n < i \leq a} \min\{I, L^{\rho_i} \|\Phi_{\alpha, \theta}^r(x) - x\|_i\}.$$

Let us observe that the conditions $\epsilon_{r, L} < \epsilon < I$ and $\delta' < \delta_{r, L}(x) < \delta < I$ are equivalent to saying that

$$\Phi_{\alpha, \theta}^r(x) = x \exp\left(\sum_{i=1}^a s_i Y_i\right)$$

for some vector $\mathbf{s} = (s_1, \dots, s_a) \in [-I/2, I/2]^a$ such that

$$\begin{aligned} |s_i| &< \epsilon L^{-\rho_i}, & \text{for all } i \in \{1, \dots, n\}; \\ |s_i| &< \delta L^{-\rho_i}, & \text{for all } i \in \{n+1, \dots, a\}; \\ |s_j| &> \delta' L^{-\rho_j} & \text{for some } j \in \{n+1, \dots, a\}. \end{aligned}$$

For every $r \in \mathbb{Z} \setminus \{0\}$ and $j \geq 0$, let $\text{AP}_{j,L}^r \subset M$ be the sets defined as follows

$$(78) \quad \text{AP}_{j,L}^r := \begin{cases} \emptyset, & \text{if } \epsilon_{r,L} > I/2; \\ (\delta_{r,L})^{-1} \left(]2^{-(j+1)}I, 2^{-j}I] \right), & \text{otherwise.} \end{cases}$$

Lemma 5.4. *For all $r \in \mathbb{Z} \setminus \{0\}$, for all $j \in \mathbb{N}$, and for all $L \geq 1$, the $(a+1)$ -dimensional Lebesgue measure \mathcal{L}^{a+1} of the set $\text{AP}_{j,L}^r$ can be estimated as follows:*

$$\mathcal{L}^{a+1}(\text{AP}_{j,L}^r) \leq \frac{I(Y)^{a-n}}{2^{j(a-n)}} L^{-\sum_{i=n+1}^a \rho_i}.$$

Proof. Without loss of generality we can assume that $\text{AP}_{j,L}^r \neq \emptyset$, otherwise there is nothing to prove. By Tonelli's Theorem,

$$(79) \quad \mathcal{L}^{a+1}(\text{AP}_{j,L}^r) = \int_0^1 \mathcal{L}_\theta^a(\text{AP}_{j,L}^r \cap \mathbb{T}_\theta^a) d\theta,$$

hence the statement can be reduced to estimates on the a -dimensional Lebesgue measure $\mathcal{L}_\theta^a(\text{AP}_{j,L}^r)$ for $\theta \in \mathbb{T}^1$. For every $j \in \mathbb{N}$, let

$$\mathcal{U}_\theta^{L,j} := \left\{ x \in \mathbb{T}_\theta^a \mid \max_{n+1 \leq i \leq a} L^{\rho_i} \|x\|_i \leq I/2^j \right\}.$$

By definition, if $x \in \text{AP}_{j,L}^r \cap \mathbb{T}_\theta^a$, then $\|\Phi_{\alpha,\theta}^r(x) - x\|_i \leq 2^{-j} I L^{-\rho_i}$ for all $i = n+1, \dots, a$, that is, $\text{AP}_{j,L}^r \cap \mathbb{T}_\theta^a \subset \text{AP}^r(\mathcal{U}_\theta^{L,j})$. By Lemma 5.2 we have

$$\begin{aligned} \mathcal{L}_\theta^a(\text{AP}_{j,L}^r \cap \mathbb{T}_\theta^a) &\leq \mathcal{L}_\theta^a(\text{AP}^r(\mathcal{U}_\theta^{L,j})) \\ &= \mathcal{L}_{\theta,\alpha,r}^{a-n}(\mathcal{U}_\theta^{L,j} \cap \mathbb{T}_{\theta,\alpha,r}^{a-n}) = \frac{I^{a-n}}{2^{j(a-n)}} L^{-\sum_{i=n+1}^a \rho_i}. \end{aligned}$$

The statement thus follows from Tonelli theorem (see formula (79)). \square

5.2. Expected width bounds. In this section we prove a bound on the average width of a quasi-Abelian nilpotent orbit with respect to a rescaled basis in terms of the ergodic average along the orbit of an appropriate function on the nilmanifold (which depends on the length of the orbit and on the rescaling exponents).

The expected value of the average width is thus bounded in terms of the average of such a function over the nilmanifold. Such an estimate is then reduced to a Diophantine estimate.

For $L \geq 1$, $r \in \mathbb{Z} \setminus \{0\}$, let us consider the function

$$(80) \quad h_{r,L} := \sum_{j=1}^{+\infty} \min\{2^{j(a-n)}, (\frac{2}{\epsilon_{r,L}})^n\} \chi_{\text{AP}_{j,L}^r}.$$

Let us introduce the cut-off $J_{r,L} \in \mathbb{N}$ by the formula:

$$(81) \quad J_{r,L} := \max\{j \in \mathbb{N} \mid 2^{j(a-n)} \leq (\frac{2}{\epsilon_{r,L}})^n\}.$$

The function in formula (80) can also be written as follows:

$$(82) \quad h_{r,L} := \sum_{j=1}^{J_{r,L}} 2^{j(a-n)} \chi_{\text{AP}_{j,L}^r} + \sum_{j > J_{r,L}} (\frac{2}{\epsilon_{r,L}})^n \chi_{\text{AP}_{j,L}^r}.$$

For every $L \geq 1$ let $\mathcal{F}_\alpha^{(L)}$ be the rescaled strongly adapted normalised basis

$$(83) \quad \begin{aligned} \mathcal{F}_\alpha^{(L)} &= (X_\alpha^{(L)}, Y_1^{(L)}, \dots, Y_a^{(L)}) \\ &= (L X_\alpha, L^{-\rho_1} Y_1, \dots, L^{-\rho_a} Y_a). \end{aligned}$$

For $(x, T) \in M \times \mathbb{R}^+$, let $w_{\mathcal{F}_\alpha^{(L)}}(x, T)$ denote the average width of the orbit segment

$$\gamma_{X_\alpha^{(L)}}^T(x) := \{\phi_{X_\alpha^{(L)}}^t(x) \mid 0 \leq t \leq T\}.$$

We prove below a bound for the average width $w_{\mathcal{F}_\alpha^{(L)}}(x, T)$ of the orbit arc $\gamma_{X_\alpha^{(L)}}^T(x)$ in terms of the following function:

$$(84) \quad H_L^T := 1 + \sum_{|r|=1}^{[TL]} h_{r,L}.$$

Lemma 5.5. *Let $\mathcal{F}_\alpha = (X_\alpha, Y)$ be any normalised strongly adapted basis. For all $x \in M$ and for all $T, L \geq 1$ we have*

$$\frac{1}{w_{\mathcal{F}_\alpha^{(L)}}(x, T)} \leq \left(\frac{2}{I(Y)} \right)^a \frac{1}{T} \int_0^T H_L^T \circ \phi_{X_\alpha^{(L)}}^t(x) dt.$$

Proof. Let x, T, L and L be defined as in the statement. For every $t \in [0, T]$, we define a set $\Omega(t) \subset \{t\} \times \mathbb{R}^a$ as follows:

- (A) If $\phi_{X_\alpha^{(L)}}^t(x) \notin \bigcup_{|r|=1}^{[TL]} \bigcup_{j>0} \text{AP}_{j,L}^r$, let $\Omega(t)$ be the set of all points (t, s_1, \dots, s_a) such that

$$|s_1| < I/4, \dots, |s_a| < I/4.$$

Observe, incidentally, that if $\phi_{X_\alpha^{(L)}}^t(x) \notin \bigcup_{|r|=1}^{[TL]} \bigcup_{j>0} \text{AP}_{j,L}^r$, then for all $|r| \in [1, TL]$ such that $\epsilon_{r,L} \leq I/2$ we must have that $x \in \text{AP}_{0,L}^r$; in fact for such an r we have $\bigcup_{j \geq 0} \text{AP}_{j,L}^r = M$.

To define the set $\Omega(t)$ when $\phi_{X_\alpha^{(L)}}^t(x) \in \bigcup_{|r|=1}^{[TL]} \bigcup_{j>0} \text{AP}_{j,L}^r$ we consider two sub-cases.

- (B) if $\phi_{X_\alpha^{(L)}}^t(x) \in \bigcup_{|r|=1}^{[TL]} \bigcup_{j>J_L^r} \text{AP}_{j,L}^r$, let $\Omega(t)$ be the set of all points (t, s_1, \dots, s_a) such that

$$|s_i| < \frac{1}{4} \min_{1 \leq |r| \leq [TL]} \min_{j > J(|r|)} \{\epsilon_{r,L} : \phi_{X_\alpha^{(L)}}^t(x) \in \text{AP}_{j,L}^r\}, \quad \text{for } i \in \{1, \dots, n\},$$

$$|s_i| < \frac{I}{4}, \quad \text{for } i \in \{n+1, \dots, a\};$$

- (C) finally, if $\phi_{X_\alpha^{(L)}}^t(x) \in \bigcup_{|r|=1}^{[TL]} \bigcup_{j \leq J_L^r} \text{AP}_{j,L}^r \setminus \bigcup_{|r|=1}^{[TL]} \bigcup_{j > J_L^r} \text{AP}_{j,L}^r$, let $\Omega(t)$, let ℓ be the largest integer such that

$$\phi_{X_\alpha^{(L)}}^t(x) \in \bigcup_{|r|=1}^{[TL]} \bigcup_{\ell \leq j \leq J_L^r} \text{AP}_{j,L}^r \setminus \bigcup_{|r|=1}^{[TL]} \bigcup_{j > J_L^r} \text{AP}_{j,L}^r$$

and let $\Omega(t)$ be the set of all points (t, s_1, \dots, s_a) such that

$$|s_i| < \frac{I}{4}, \quad \text{for } i \in \{1, \dots, n\}$$

$$|s_i| < \frac{I}{4} \frac{1}{2^{\ell+1}}, \quad \text{for } i \in \{n+1, \dots, a\};$$

Then we set

$$\Omega := \bigcup_{t \in [0, T]} \Omega(t) \subset [0, T] \times \mathbb{R}^a.$$

It is clear that the set Ω contains the segment $[0, T] \times \{0\}$ and it is contained in $[0, T] \times [-4^{-1}I, 4^{-1}I]^a$.

We claim that the restriction to Ω of the map

$$(85) \quad (t, s_1, \dots, s_a) \in \Omega \mapsto x \exp(tX_\alpha^{(L)}) \exp(s_1 Y_1^{(L)} + \dots + s_a Y_a^{(L)}) \in M$$

is injective. In fact, let us assume that there exist points $(t, s_1, \dots, s_a) \in \Omega$ and $(t', s'_1, \dots, s'_a) \in \Omega$ such that

$$(86) \quad \begin{aligned} & \phi_{X_\alpha^{(L)}}^t(x) \exp(s_1 Y_1^{(L)}) + \dots + s_a Y_a^{(L)} \\ &= \phi_{X_\alpha^{(L)}}^{t'}(x) \exp(s'_1 Y_1^{(L)} + \dots + s'_a Y_a^{(L)}). \end{aligned}$$

We can assume $t' \geq t$. By considering the projection on the base torus \mathbb{T}^{n+1} we have the following identity:

$$(87) \quad \begin{aligned} (t, s_1, \dots, s_n) \mod \mathbb{Z}^{n+1} &= \text{pr}_1(\phi_{X_\alpha^{(L)}}^t(x)) \\ &= \text{pr}_1(\phi_{X_\alpha^{(L)}}^{t'}(x)) = (t', s'_1, \dots, s'_n) \mod \mathbb{Z}^{n+1}; \end{aligned}$$

this implies $t \equiv t'$ modulo \mathbb{Z} . As $\phi_{X_\alpha^{(L)}}^t = \phi_{X_\alpha}^{tL}$, the number $r_0 = t' - t$ is a non negative integer satisfying $r_0 \leq TL$; hence $r_0 \leq [TL]$.

If $r_0 = 0$, then $t = t'$ and $s_1 = s'_1, \dots, s_a = s'_a$: in fact, by the definition of the constant I , the map

$$(s_1, \dots, s_a) \in [-4^{-1}I, 4^{-1}I]^a \mapsto \phi_{X_\alpha^{(L)}}^t(x) \exp(s_1 Y_1^{(L)}) \dots \exp(s_a Y_a^{(L)})$$

is injective. We prove below that the overlapping identity (86) leads to a contradiction if we assume that $r_0 \neq 0$.

The condition (87) tells us that the points

$$p := \phi_{X_\alpha^{(L)}}^t(x) \quad \text{and} \quad q := \phi_{X_\alpha^{(L)}}^{t'}(x)$$

belong to the same torus \mathbb{T}_θ^a ; the definition of r_0 says that $q = \Phi_{\alpha, \theta}^{r_0}(p)$. From identity (86) we have

$$\begin{aligned} q &= p \exp((s'_1 - s_1)Y_1^{(L)} + \dots + (s'_a - s_a)Y_a^{(L)}) \\ &= p \exp((s'_1 - s_1)L^{-\rho_1}Y_1 + (s'_2 - s_2)L^{-\rho_2}Y_2 \dots + (s'_a - s_a)L^{-\rho_a}Y_a); \end{aligned}$$

thus, for all $i \in \{1, \dots, a\}$,

$$(88) \quad \begin{aligned} L^{\rho_i} \|p - \Phi_{\alpha, \theta}^{r_0}(p)\|_i &= L^{\rho_i} \|q - \Phi_{\alpha, \theta}^{-r_0}(q)\|_i \\ &\leq L^{\rho_i} |(s'_i - s_i)L^{-\rho_i}| \leq |s_i| + |s'_i|. \end{aligned}$$

In particular, from formula (88) we obtain that $\epsilon_{r_0, L}$, which is a constant on M , satisfies the following inequality:

$$(89) \quad \epsilon_{r_0, L} = \max_{1 \leq i \leq n} L^{\rho_i} \|p - \Phi_{\alpha, \theta}^{r_0}(p)\|_i \leq \max_{1 \leq i \leq n} |s_i| + |s'_i| < I/2.$$

For the same reason, that is, from formula (88), we also obtain that

$$\delta_{r_0, L}(p) = \delta_{-r_0, L}(q) < I/2.$$

By defining $j_0 \in \mathbb{N}$ as the unique non-negative integer such that

$$(90) \quad \frac{I}{2^{j_0+1}} < \delta_{r_0, L}(p) \leq \frac{I}{2^{j_0}},$$

and by the definition (78), we have that $p \in \text{AP}_{j_0, L}^{r_0}$ and $q \in \text{AP}_{j_0, L}^{-r_0}$.

If $j_0 > J_L^{r_0} = J_L^{-r_0}$, then $p, q \in \bigcup_{0 < |r| \leq [TL]} \bigcup_{j > J_L^r} \text{AP}_{j, L}^r$; it follows that the sets $\Omega(t)$ and $\Omega(t')$ are both defined according to definition (B); hence, from (89) and the definition (B), we obtain

$$\epsilon_{r_0, L} \leq \max_{1 \leq i \leq n} |s_i| + |s'_i| \leq \frac{1}{2} \epsilon_{r_0, L},$$

a plain contradiction.

Our conclusion at this point is that if the map in formula (85) fails to be injective at the points (t, s_1, \dots, s_a) , (t', s'_1, \dots, s'_a) , with $t \leq t'$, then there are integers $r_0 \in [1, TL]$, $j_0 \in [1, J(|r_0|)]$ and $\theta \in \mathbb{T}^1$ such that the points $p = x \exp(tX_\alpha^{(L)})$, $q = x \exp(tX_\alpha^{(L)})$ satisfy

$$p, q \in \mathbb{T}_\theta^a, \quad q = \Psi_{\alpha, \theta}^{r_0}(p), \quad p \in \text{AP}_{j_0, L}^{r_0}, \quad q \in \text{AP}_{j_0, L}^{-r_0};$$

in addition,

$$p, q \notin \bigcup_{0 < |r| \leq [TL]} \bigcup_{j > J_L^r} \text{AP}_{j, L}^r.$$

In this case the sets $\Omega(t)$ and $\Omega(t')$ are both defined according to definition (C); by defining ℓ_1 and ℓ_2 as the largest integers such that

$$p \in \bigcup_{0 < |r| \leq [TL]} \bigcup_{\ell_1 \leq j \leq J_L^r} \text{AP}_{j, L}^r$$

and

$$q \in \bigcup_{0 < |r| \leq [TL]} \bigcup_{\ell_2 \leq j \leq J_L^r} \text{AP}_{j, L}^r,$$

we have, from definition (C),

$$|s_i| < \frac{I}{4} \frac{1}{2^{\ell_1+1}}, \quad |s'_i| < \frac{I}{4} \frac{1}{2^{\ell_2+1}}, \quad \text{for all } i \in \{n+1, \dots, a\};$$

this also leads to a contradiction because from formulas (88), (90), since by construction $\ell_1, \ell_2 \geq j_0$, we deduce that

$$\frac{I}{2^{j_0+1}} \leq \delta_{r_0, L}(p) \leq \max_{i \in \{n+1, \dots, a\}} \{|s_i| + |s'_i|\} < \frac{I}{4} \frac{1}{2^{\ell_1+1}} + \frac{I}{4} \frac{1}{2^{\ell_2+1}} \leq \frac{I}{2} \frac{1}{2^{j_0+1}},$$

again a plain contradiction. The injectivity claim is therefore proved.

We are finally ready to conclude the proof. In fact the width function w_Ω of the set Ω is given, by definition, by the following formulas:

$$w_\Omega(t) = \begin{cases} \left(\frac{I}{2}\right)^a & \text{if } \phi_{X_\alpha}^t(x) \notin \bigcup_{|r|=1}^{[TL]} \bigcup_{j>0} \text{AP}_{j,L}^r; \\ \left(\frac{I}{2}\right)^{a-n} \left(\frac{\min\{\epsilon_{r,L}\}}{2}\right)^n & \text{if } \phi_{X_\alpha}^t(x) \in \bigcup_{|r|=1}^{[TL]} \bigcup_{j>J_L^r} \text{AP}_{j,L}^r \text{ and} \\ & \text{where the min is over all } |r| \in [1, TL] \\ & \text{such that } \phi_{X_\alpha}^t(x) \in \bigcup_{j>J_L^r} \text{AP}_{j,L}^r; \\ \left(\frac{I}{2}\right)^a 2^{-(a-n)(\ell+1)} & \text{in the remaining case, where } \ell \text{ is the} \\ & \text{largest integer } \leq \max\{J_L^r | r \in [-TL, TL]\} \\ & \text{such that } \phi_{X_\alpha}^t(x) \in \bigcup_{|r|=1}^{[TL]} \bigcup_{j=\ell}^{J_L^r} \text{AP}_{j,L}^r. \end{cases}$$

In the first case we have

$$\frac{1}{w_\Omega(t)} \leq \left(\frac{2}{I}\right)^a;$$

in the second case

$$\frac{1}{w_\Omega(t)} \leq \left(\frac{2}{I}\right)^{a-n} \sum_{|r|=1}^{[TL]} \sum_{j>J_L^r} \frac{2^n \chi_{\text{AP}_{j,L}^r}(\phi_{X_\alpha}^t(x))}{(\epsilon_{r,L})^n};$$

in the third and last case

$$\frac{1}{w_\Omega(t)} \leq \left(\frac{2}{I}\right)^{a-n} \sum_{|r|=1}^{[TL]} \sum_{j=1}^{J_L^r} 2^{(a-n)(j+1)} \chi_{\text{AP}_{j,L}^r}(\phi_{X_\alpha}^t(x)).$$

Thus, by the definition of the function H_L^T in formula (84), we have

$$\frac{1}{w_\Omega(t)} \leq \left(\frac{2}{I}\right)^a H_L^T \circ \phi_{X_\alpha}^t(x), \quad \text{for all } t \in [0, T].$$

From the definition (28) of the average width $w_{\mathcal{F}_\alpha^{(L)}}(x, T)$ of the orbit segment $\{x \exp(tX_\alpha^{(L)}) | 0 \leq t \leq T\}$, we have the estimate

$$\frac{1}{w_{\mathcal{F}_\alpha^{(L)}}(x, T)} \leq \frac{1}{T} \int_0^T \frac{dt}{w_\Omega(t)} \leq \left(\frac{2}{I}\right)^a \frac{1}{T} \int_0^T H_L^T \circ \phi_{X_\alpha}^t(x) dt.$$

The argument is therefore completed. \square

By Lemma 5.5, a bound on the expected value of the inverse of the average width can be derived from the following integral estimate:

Lemma 5.6. *For all $r \in \mathbb{Z} \setminus \{0\}$ and for all $L \geq 1$ the following estimate holds:*

$$\left| \int_M h_{r,L}(x) dx \right| \leq I(Y)^{a-n} (1 + J_L^r) L^{-\sum_{i=n+1}^a \rho_i}.$$

Proof. By Lemma 5.4 it follows that, for all $r \neq 0$ and for all $j \geq 0$, the Lebesgue measure of the set $\text{AP}_{j,L}^r$ satisfies the following bound:

$$(91) \quad \mathcal{L}^{a+1}(\text{AP}_{j,L}^r) \leq \frac{I^{a-n}}{2^{j(a-n)}} L^{-\sum_{i=n+1}^a \rho_i}.$$

We are now ready to estimate the integral of the function $h_{r,L}$; in fact, by formula (82), it follows that

$$\begin{aligned} \int_M h_{r,L}(x) dx &\leq 1 + \sum_{i=1}^{J_L^r} 2^{j(a-n)} \mathcal{L}^{a+1}(\text{AP}_{j,L}^r) \\ &\quad + \sum_{j>J_L^r} \frac{2^n \mathcal{L}^{a+1}(\text{AP}_{j,L}^r)}{(\epsilon_{r,L})^n}; \end{aligned}$$

by the estimate in formula (91), we immediately have that

$$\sum_{i=1}^{J_L^r} 2^{j(a-n)} \mathcal{L}^{a+1}(\text{AP}_{j,L}^r) \leq I^{a-n} J_L^r L^{-\sum_{i=n+1}^a \rho_i}.$$

By the definition of the cut-off in formula (81) we have the bound

$$\frac{2^{n-(J_L^r+1)(a-n)}}{(\epsilon_{r,L})^n} \leq 1,$$

and, by an elementary estimate on a geometric sum,

$$\sum_{j>J_L^r} \frac{2^n \mathcal{L}^{a+1}(\text{AP}_{j,L}^r)}{(\epsilon_{r,L})^n} \leq \frac{2^{n-(J_L^r+1)(a-n)}}{(\epsilon_{r,L})^n} I^{a-n} L^{-\sum_{i=n+1}^a \rho_i},$$

hence the statement follows. \square

5.3. Diophantine estimates. In this section we state the relevant Diophantine condition for simultaneous Diophantine approximations in any dimensions. We then derive bounds on the expected average width under Diophantine conditions from the results of the previous section. We also prove that in dimension one our Diophantine condition is equivalent to the standard Diophantine condition.

Definition 5.7. For any basis $\bar{Y} := \{\bar{Y}_1, \dots, \bar{Y}_n\} \subset \mathbb{R}^n$, let $\bar{I} := \bar{I}(\bar{Y})$ be the supremum of all constants $\bar{I}' > 0$ such that the map

$$(s_1, \dots, s_n) \rightarrow \exp\left(\sum_{i=1}^n s_i \bar{Y}_i\right) \in \mathbb{T}^n.$$

is a local embedding (injective) on the domain

$$\{\mathbf{s} \in \mathbb{R}^n \mid |s_i| < \bar{I}' \text{ for all } i = 1, \dots, n\}.$$

For any $\theta \in \mathbb{R}^n$, let $[\theta] \in \mathbb{T}^n$ its projection onto the torus $\mathbb{T}^n := \mathbb{R}^n / \mathbb{Z}^n$ and let

$$|\theta|_1 = |s_1|, \dots, |\theta|_i = |s_i|, \dots, |\theta|_n = |s_n|,$$

if there is $\mathbf{s} := (s_1, \dots, s_n) \in [-\bar{I}/2, \bar{I}/2]^n$ such that

$$[\theta] = \exp\left(\sum_{i=1}^n s_i \bar{Y}_i\right) \in \mathbb{T}^n;$$

otherwise we set $|\theta|_1 = \dots = |\theta|_n = \bar{I}$.

We introduce below our (simultaneous) Diophantine condition in all dimensions.

Definition 5.8. Let $\sigma := (\sigma_1, \dots, \sigma_n) \in (0, 1)^n$ be such that $\sigma_1 + \dots + \sigma_n = 1$. For any $\alpha := (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$, for any $N \in \mathbb{N}$ and for every $\delta > 0$, let

$$\mathcal{R}_\alpha^{(\bar{Y}, \sigma)}(N, \delta) = \{r \in [-N, N] \cap \mathbb{Z} \mid |r\alpha|_1 \leq \delta^{\sigma_1}, \dots, |r\alpha|_n \leq \delta^{\sigma_n}\}.$$

For every $\nu \geq 1$, let $D_n(\bar{Y}, \sigma, \nu) \subset (\mathbb{R} \setminus \mathbb{Q})^n$ be the subset defined as follows: the vector $\alpha \in D_n(\bar{Y}, \sigma, \nu)$ if and only if there exists a constant $C(\bar{Y}, \sigma, \alpha) > 0$ such that, for all $N \in \mathbb{N}$ and for all $\delta > 0$,

$$(92) \quad \#\mathcal{R}_\alpha^{(\bar{Y}, \sigma)}(N, \delta) \leq C(\bar{Y}, \sigma, \alpha) \max\{N^{1-\frac{1}{\nu}}, N\delta\}.$$

Let us prove that in all dimensions the above Diophantine condition implies a standard simultaneous Diophantine condition.

Lemma 5.9. *Let $\alpha \in D_n(\bar{Y}, \sigma, \nu)$. For all $r \in \mathbb{Z} \setminus \{0\}$, we have*

$$\max\{|r\alpha|_1, \dots, |r\alpha|_n\} \geq \min\left\{\frac{\bar{I}^2}{4}, \frac{1}{[1 + C(\bar{Y}, \sigma, \alpha)]^{2\nu}}\right\} \frac{1}{|r|^\nu}.$$

Proof. Let $C := \max\{2/\bar{I}, [1 + C(\bar{Y}, \sigma, \alpha)]^\nu\}$. Since $\nu \geq 1$, we have

$$(93) \quad C \geq 1 + C(\bar{Y}, \sigma, \alpha)^\nu \geq 1 + C(\bar{Y}, \sigma, \alpha) > 1.$$

Let us assume by contradiction that there exists $r \in \mathbb{Z} \setminus \{0\}$ such that

$$\max\{|r\alpha|_1, \dots, |r\alpha|_n\} < \frac{1}{C^2|r|^\nu},$$

For all $k \in \{|r|, 2|r|, \dots, [C|r|^{\nu-1}] \times |r|\}$, we have $\frac{k}{C^2|r|^\nu} \leq \bar{I}/2$. It follows that

$$|k\alpha|_i < \frac{1}{C|r|} \leq \left(\frac{1}{C|r|}\right)^{\sigma_i}, \quad \text{for all } i \in \{1, \dots, n\}.$$

hence by the definitions (in particular, by the estimate in formula (92))

$$\begin{aligned} [C|r|^{\nu-1}] &\leq \#\mathcal{R}_\alpha^{(\bar{Y}, \sigma)}([C|r|^{\nu-1}] \times |r|, \frac{1}{C|r|}) \\ &\leq C(\bar{Y}, \sigma, \alpha) \max\{[C|r|^{\nu-1}]^{1-\frac{1}{\nu}} \times |r|^{1-\frac{1}{\nu}}, [C|r|^{\nu-1}]/C\}. \end{aligned}$$

Since $C > C(\bar{Y}, \sigma, \alpha)$ by formula (93), from the above inequality we derive

$$[C|r|^{\nu-1}] \leq C(\bar{Y}, \sigma, \alpha)^\nu \times |r|^{\nu-1}.$$

which, by taking into account that $|r|^{\nu-1} \geq 1$, implies that $C < 1 + C(\bar{Y}, \sigma, \alpha)^\nu$, in contradiction with the inequality in formula (93). \square

We prove below that the Diophantine condition introduced above in Definition 5.8 follows from a standard simultaneous Diophantine condition (of different exponent). This results implies that for any $\nu > 1$ our condition holds for a full measure set of vectors. In dimension one our condition coincides with the classical Diophantine condition of the same exponent. The proof in dimension one is an exercise based on continued fractions. We owe the proof in the general case, which we explain below, to a personal communication of N. Chevallier.

Lemma 5.10. *Let $\{(q_i, p_i)\} \subset \mathbb{N} \times \mathbb{Z}^n$ denote the sequence of best approximation vectors of a vector $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n \setminus \mathbb{Q}^n$ with respect to the sup norm $\|\cdot\|$ (or to any other norm). For all $i \in \mathbb{N}$, we adopt the notation*

$$\epsilon_i := q_i\alpha - p_i \in \mathbb{R}^n \quad \text{and} \quad d_i := d(q_i\alpha, \mathbb{Z}^n) = \|q_i\alpha - p_i\| > 0.$$

For any vector $\sigma = (\sigma_1, \dots, \sigma_n) \in (0, 1)^n$ such that $\sigma_1 + \dots + \sigma_n = 1$, let

$$m(\sigma) := \min\{\sigma_1, \dots, \sigma_n\} \quad \text{and} \quad M(\sigma) := \max\{\sigma_1, \dots, \sigma_n\}.$$

The vector $\alpha \in D_n(\bar{Y}, \sigma, \nu)$ for any basis $\bar{Y} \subset \mathbb{R}^n$ and for any $\nu \geq 1$ under the assumption that there exists a constant $C_\alpha > 0$ such that, for all $i \in \mathbb{N}$,

$$\begin{aligned} (a) \quad q_{i+1} &\leq C_\alpha q_i^\nu, \quad (b) \quad q_{i+1}^{M(\sigma)/\nu} \leq C_\alpha q_i d_{i-1}^{n-1}, \\ (c) \quad q_{i+1}^{1/\nu} &\leq C_\alpha q_i d_{i-1}^{(n-2)(1-\frac{m(\sigma)}{M(\sigma)})}. \end{aligned}$$

Proof. Let $N \in \mathbb{N}$ be a positive integer such that $q_i \leq N < q_{i+1}$. For any $r \in [-N, N] \cap \mathbb{Z}$, by the Euclidean algorithm, there exist $a \in \mathbb{Z}$ and $0 \leq b < q_i$ such that $r = a q_i + b$. It follows that $r\alpha = b\alpha + a\epsilon_i$ modulo \mathbb{Z}^n , hence by definition we have that

$$(94) \quad |r\alpha|_m = |b\alpha + a\epsilon_i|_m, \quad \text{for all } m \in \{1, \dots, n\}.$$

For any $\theta \in \mathbb{T}^n$ and $\delta \in (0, 1)$, let $N_i^{(\sigma)}(\theta, \delta)$ be defined as follows:

$$N_i^{(\sigma)}(\theta, \delta) := \#\{b \in [0, q_i - 1] \cap \mathbb{N} \mid |b\alpha + \theta|_1 \leq \delta^{\sigma_1}, \dots, |b\alpha + \theta|_n \leq \delta^{\sigma_n}\}.$$

By formula (94), it follows that

$$(95) \quad \#\mathcal{R}_\alpha^{(\bar{Y}, \sigma)}(N, \delta) \leq \sum_{a=0}^{\lfloor N/q_i \rfloor} N_i^{(\sigma)}(a\epsilon_i, \delta).$$

We are therefore led to estimate the integers $N_i^{(\sigma)}(\theta, \delta)$ for any point $\theta \in \mathbb{T}^n$.

For all $i \in \mathbb{N}$, let $\lambda_{i,1} \leq \dots \leq \lambda_{i,n}$ denote the minima of the lattice

$$\Lambda_i := \mathbb{Z} \frac{p_i}{q_i} + \mathbb{Z}^n.$$

Let us remark that by the definition of the best approximation vectors it follows that the first minimum $\lambda_{i,1}$ satisfies the estimate

$$\lambda_{i,1} \leq 2d_{i-1}.$$

Let $(\tau_1, \dots, \tau_n) \in (0, 1)^n$ be a permutation of $\{\sigma_1, \dots, \sigma_n\}$ such that

$$\tau_1 \leq \dots \leq \tau_n.$$

There exists a constant $C_n(\bar{Y}) > 0$ such that, for all $\theta \in \mathbb{T}^n$, we have

$$\begin{aligned} &\#\{z \in \Lambda_i \mid |z + \theta|_1 \leq \delta^{\sigma_1}, \dots, |z + \theta|_n \leq \delta^{\sigma_n}\} \\ &\leq C_n(\bar{Y}) \left(1 + \frac{\delta^{\tau_1}}{\lambda_{i,1}} + \frac{\delta^{\tau_1+\tau_2}}{\lambda_{i,1}\lambda_{i,2}} \dots + \frac{\delta}{\lambda_{i,1} \dots \lambda_{i,n}} \right) \\ &\leq C_n(\bar{Y}) n \left\{ 1 + \max\left[\frac{\delta^{\tau_1}}{d_{i-1}}, \dots, \frac{\delta^{\tau_1+\dots+\tau_{n-1}}}{d_{i-1}^{n-1}}, \frac{\delta}{\det \Lambda_n} \right] \right\}. \end{aligned}$$

By taking into account that $d(b\alpha, bp_i/q_i) \leq d_i \leq d_{i-1}$, for all $b \in [0, q_i - 1] \cap \mathbb{N}$, it follows from the above formula that there exists a constant $C'_n(\bar{Y}) > 0$ such that, for any $\theta \in \mathbb{T}^n$ and $\delta \in (0, 1)$ we have

$$N_i^{(\sigma)}(\theta, \delta) \leq C'_n(\bar{Y}) \left\{ 1 + \max\left[\frac{\delta^{\tau_1}}{d_{i-1}}, \dots, \frac{\delta^{\tau_1+\dots+\tau_{n-1}}}{d_{i-1}^{n-1}}, \delta q_i \right] \right\},$$

hence by formula (95), whenever $q_i \leq N < q_{i+1}$ and for all $\delta \in (0, 1)$ we have

$$\#\mathcal{R}_\alpha^{(\bar{Y}, \sigma)}(N, \delta) \leq C'_n(\bar{Y}) \frac{N}{q_i} \{1 + \max[\frac{\delta^{\tau_1}}{d_{i-1}^{n-1}}, \dots, \frac{\delta^{\tau_1 + \dots + \tau_{n-1}}}{d_{i-1}^{n-1}}, \delta q_i]\}.$$

We distinguish two cases:

$$\begin{aligned} (1) \quad & \max[\frac{\delta^{\tau_1}}{d_{i-1}^{n-1}}, \dots, \frac{\delta^{\tau_1 + \dots + \tau_{n-1}}}{d_{i-1}^{n-1}}] \leq \delta q_i; \\ (2) \quad & \max[\frac{\delta^{\tau_1}}{d_{i-1}^{n-1}}, \dots, \frac{\delta^{\tau_1 + \dots + \tau_{n-1}}}{d_{i-1}^{n-1}}] > \delta q_i. \end{aligned}$$

In case (1), by taking into account that by hypothesis $C_\alpha^{1/\nu} q_i \geq q_{i+1}^{1/\nu} \geq N^{1/\nu}$, we derive the following upper bound:

$$(96) \quad \#\mathcal{R}_\alpha^{(\bar{Y}, \sigma)}(N, \delta) \leq C'_n(\bar{Y}) \{ \frac{N}{q_i} + N\delta \} \leq C'_n(\bar{Y}) \{ C_\alpha^{1/\nu} N^{1-1/\nu} + N\delta \}.$$

In case (2) we distinguish two sub-cases:

$$\begin{aligned} (2a) \quad & \max[\frac{\delta^{\tau_1}}{d_{i-1}^{n-1}}, \dots, \frac{\delta^{\tau_1 + \dots + \tau_{n-1}}}{d_{i-1}^{n-1}}] = \frac{\delta^{\tau_1 + \dots + \tau_{n-1}}}{d_{i-1}^{n-1}}, \\ (2b) \quad & \max[\frac{\delta^{\tau_1}}{d_{i-1}^{n-1}}, \dots, \frac{\delta^{\tau_1 + \dots + \tau_{n-1}}}{d_{i-1}^{n-1}}] \neq \frac{\delta^{\tau_1 + \dots + \tau_{n-1}}}{d_{i-1}^{n-1}}. \end{aligned}$$

In case (2a) we have

$$\delta q_i < \frac{\delta^{\tau_1 + \dots + \tau_{n-1}}}{d_{i-1}^{n-1}} = \frac{\delta^{1-M(\sigma)}}{d_{i-1}^{n-1}},$$

so that by our assumption (b) we can derive the following upper bound:

$$\frac{\delta^{\tau_1 + \dots + \tau_{n-1}}}{q_i d_{i-1}^{n-1}} \leq \left(\frac{1}{q_i d_{i-1}^{n-1}} \right)^{\frac{1}{M(\sigma)}} \leq C_\alpha^{1/\nu} q_{i+1}^{-1/\nu} \leq C_\alpha^{1/\nu} N^{-1/\nu}.$$

In case (2b), let $j < n-1$ be such that

$$\max[\frac{\delta^{\tau_1}}{d_{i-1}^{n-1}}, \dots, \frac{\delta^{\tau_1 + \dots + \tau_{n-1}}}{d_{i-1}^{n-1}}] = \frac{\delta^{\tau_1 + \dots + \tau_j}}{d_{i-1}^j}.$$

Since the above condition immediately implies that $\delta^{\tau_{j+1}} \leq d_{i-j}$, it follows from our assumption (c) that the following upper bound holds:

$$\frac{\delta^{\tau_1 + \dots + \tau_j}}{q_i d_{i-1}^j} \leq \frac{1}{q_i} \left(\frac{1}{d_{i-1}^{n-2}} \right)^{1 - \frac{m(\sigma)}{M(\sigma)}} \leq C_\alpha^{1/\nu} q_{i+1}^{-1/\nu} \leq C_\alpha^{1/\nu} N^{-1/\nu}.$$

We have therefore proved that under our assumptions the upper bound in formula (96) holds also in case (2), hence $\alpha \in D_n(\bar{Y}, \sigma, \nu)$. \square

Let us recall the classical definition of a simultaneously Diophantine vector.

Definition 5.11. A vector $\alpha \in \mathbb{R}^n \setminus \mathbb{Q}^n$ is simultaneously Diophantine of exponent $\nu \geq 1$ if there exists a constant $c(\alpha) > 0$ such that, for all $r \in \mathbb{N} \setminus \{0\}$,

$$\|r\alpha\|_{\mathbb{Z}^n} \geq \frac{c(\alpha)}{r^{\nu/n}}.$$

Let $DC_{n,\nu} \subset \mathbb{R}^n \setminus \mathbb{Q}^n$ denote the set of all simultaneously Diophantine vectors of exponent $\nu \geq 1$.

Lemma 5.12. *For all bases $\bar{Y} \subset \mathbb{R}^n$, for all $\sigma = (\sigma_1, \dots, \sigma_n) \in (0, 1)^n$ such that $\sigma_1 + \dots + \sigma_n = 1$ and for all $\nu \geq 1$, the inclusion*

$$(97) \quad DC_{n,\mu} \subset D_n(\bar{Y}, \sigma, \nu).$$

holds under the assumption that

$$(98) \quad \mu \leq \min\left\{\nu, \left[\frac{M(\sigma)}{\nu} + 1 - \frac{1}{n}\right]^{-1}, \left[\frac{1}{\nu} + \left(1 - \frac{2}{n}\right)\left(1 - \frac{m(\sigma)}{M(\sigma)}\right)\right]^{-1}\right\}.$$

In particular, the set $D_n(\bar{Y}, \sigma, \nu)$ has full Lebesgue measure if

$$(99) \quad 1/\nu < \min\{[M(\sigma)n]^{-1}, 1 - (1 - \frac{2}{n})(1 - \frac{m(\sigma)}{M(\sigma)})\}.$$

Proof. The inclusion in formula (97) under the conditions in formula (98) follows from Lemma 5.10. In fact, by elementary calculations it is possible to prove, taking into account that, for all $i \in \mathbb{N}$, we always have

$$d_i \leq d_{i-1} \quad \text{and} \quad q_{i+1}d_i^m \leq 1,$$

and under the assumption that $\alpha \in DC_{n,\mu}$ we also have that, for all $i \in \mathbb{N}$,

$$d_i = \|q_i\alpha - p_i\| = \|q_i\alpha\|_{\mathbb{Z}^n} \geq c(\alpha)q_i^{-\mu/n},$$

that the following holds. Condition (a) of Lemma 5.10 holds if $\mu \leq \nu$, condition (b) of Lemma 5.10 holds if

$$\mu \leq \left[\frac{M(\sigma)}{\nu} + 1 - \frac{1}{n}\right]^{-1},$$

and, finally, condition (c) of Lemma 5.10 holds if

$$\mu \leq \left[\frac{1}{\nu} + \left(1 - \frac{2}{n}\right)\left(1 - \frac{m(\sigma)}{M(\sigma)}\right)\right]^{-1}.$$

The first part of the proof is therefore completed.

The proof of the second part is based on the classical elementary fact that the set $DC_{n,\mu} \subset \mathbb{R}^n$ has full Lebesgue measure for all $\mu > 1$. It follows that the set $D_n(\bar{Y}, \sigma, \nu)$ has full Lebesgue measure whenever the minimum on the right hand side of formula (98) is strictly larger than 1. By an elementary calculation one can prove that this condition is verified if the inequality in formula (99) holds. The proof of the second part of the statement is therefore completed as well. \square

In dimension one since the vector space \mathbb{R} has a unique basis up to scaling, the Diophantine condition introduced in Definition 5.8 above is independent of the choice of the basis $\bar{Y} \subset \mathbb{R}$ and of the probability vector $\sigma \in \mathbb{R}$. We therefore omit the pair (\bar{Y}, σ) from the notations introduced above. The following result is an immediate consequence of Lemma 5.9 and of Lemma 5.12.

Lemma 5.13. *For all $\nu \geq 1$ the following identity holds:*

$$DC_{1,\nu} = D_1(\nu).$$

We can finally proceed to derive our main bound on the expected width under the above Diophantine condition.

Let $\mathcal{F}_\alpha := (X_\alpha, Y)$ be a strongly adapted basis and let $\bar{Y} = (\bar{Y}_1, \dots, \bar{Y}_n) \in \mathbb{R}^n$ denote the projection of the basis $Y = (Y_1, \dots, Y_n)$ of the Abelian ideal $\mathfrak{a} \subset \mathfrak{g}$ onto the Abelianised Lie algebra $\bar{\mathfrak{g}} := \mathfrak{g}/[\mathfrak{g}, \mathfrak{g}] \approx \mathbb{R}^n$.

For any $\rho := (\rho_1, \dots, \rho_a) \in [0, 1]^a$, let us adopt the notation

$$\bar{\rho} := (\rho_1, \dots, \rho_n), \quad |\bar{\rho}| := \rho_1 + \dots + \rho_n.$$

Let $\alpha_1 := (\alpha_1^{(1)}, \dots, \alpha_1^{(n)}) \in D_n(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \nu)$. For brevity we also adopt the following notation. Let $C(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \alpha_1)$ denote the constant in the Diophantine condition introduced in Definition 5.8 and let

$$(100) \quad C(\alpha_1) := 1 + C(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \alpha_1).$$

We prove below an upper bound on the cut-off function introduced in formula (81). Let us recall that by definition, for all $L \geq 1$ and for all $r \in \mathbb{Z} \setminus \{0\}$, we have

$$J_L^r = \max\{j \in \mathbb{N} | 2^{j(a-n)} \leq (\frac{2}{\epsilon_{r,L}})^n\}.$$

The following logarithmic upper bound holds. Let $I = I(Y)$ and $\bar{I} = \bar{I}(\bar{Y})$ be the positive constants introduced in Definition 5.3 and Definition 5.7. We observe that by definition $I(Y) \leq \bar{I}(\bar{Y})$ since the basis $\bar{Y} \subset \mathbb{R}^n$ is the projection of the basis $Y \subset \mathfrak{a}$ and the canonical projection commutes with the exponential maps.

Lemma 5.14. *For every $\rho \in [0, 1]^a$, for every $\nu \leq 1/|\bar{\rho}|$ and for every $\alpha_1 := (\alpha_1^{(1)}, \dots, \alpha_1^{(n)}) \in D_n(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \nu)$ there exists a constant $K := K(a, n, \nu) > 0$ such that, for all $L \geq 1$ and for all $r \in \mathbb{Z} \setminus \{0\}$, the following bound holds:*

$$J_L^r \leq K\{1 + \log^+[I(Y)^{-1}] + \log C(\alpha_1)\}(1 + \log |r|).$$

Proof. By Lemma 5.9 and by the definition of $\epsilon_{r,L}$ in formula (77), it follows that, since $\alpha_1 \in D_n(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \nu)$, for all $T > 0$, $L \geq 1$ and for all $r \in \mathbb{Z} \setminus \{0\}$, we have

$$\epsilon_{r,L} = \max_{1 \leq i \leq n} \min\{I, |r\alpha_1|_i\} \geq \min\{I, \frac{\bar{I}^2}{4}, \frac{1}{[1 + C(\alpha_1)]^{2\nu}}\} \frac{1}{|r|^\nu}.$$

It follows by the above bound and by the definition of the cut-off that

$$J_L^r \leq \frac{n}{a-n}(3 \log 2 + 3 \log^+(1/I) + 2\nu \log[1 + C(\alpha_1)] + \nu \log |r|),$$

hence the statement follows. \square

Let $\mathcal{F}_\alpha := (X_\alpha, Y)$ be a strongly adapted basis and let $\rho \in [0, 1]^a$ a vector of scaling exponents. Assume that there exists $\nu \leq 1/|\bar{\rho}|$ such that $\alpha_1 \in DC_n(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \nu)$. For brevity, we introduce the following notation:

$$(101) \quad \mathcal{H}(Y, \rho, \alpha) := 1 + I(Y)^{a-n} C(\alpha_1) \{1 + \log^+[I(Y)^{-1}] + \log C(\alpha_1)\}.$$

The following bound holds.

Theorem 5.15. *For every $\rho \in [0, 1]^a$, for every $\nu \leq 1/|\bar{\rho}|$ and for every $\alpha_1 := (\alpha_1^{(1)}, \dots, \alpha_1^{(n)}) \in D_n(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \nu)$ there exists a constant $K' := K'(a, n, \nu) > 0$ such that, for all $T > 0$ and for all $L \geq 1$, the following bound holds:*

$$|\int_M H_L^T(x) dx| \leq K' \mathcal{H}(Y, \rho, \alpha) (1 + T) (1 + \log^+ T + \log L) L^{(1 - \sum_{i=1}^a \rho_i)}.$$

Proof. By the definition of the function H_L^T in formula (84), the statement follows from Lemma 5.6 and Lemma 5.14. In fact, for all $r \in \mathbb{Z} \setminus \{0\}$ and all $j \geq 0$, by definition (78) the set $\text{AP}_{j,L}^r$ is non-empty only if $\epsilon_{r,L} < I/2$. Since $\nu \leq 1/|\bar{\rho}|$, it follows directly from the definition of the Diophantine class $D_n(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \nu)$, that for all $T, L \geq 1$, we have

$$\#\{r \in [-TL, TL] \cap \mathbb{Z} \setminus \{0\} \mid \text{AP}_{j,L}^r \neq \emptyset\} \leq C(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \alpha_1)(1+T)L^{1-|\bar{\rho}|},$$

hence by Lemma 5.6 and Lemma 5.14 the statement is proved. \square

5.4. Width estimates along orbit segments. Let $\rho := (\rho_1, \dots, \rho_a) \in [0, 1]^a$ and let $\mathcal{F}_\alpha = (X_\alpha, Y)$ be a normalised strongly adapted basis, and recall the notation (83) for the rescaled bases $\mathcal{F}_\alpha^{(L)}$.

Definition 5.16. For any increasing sequence (L_i) of positive real numbers, let $N_i := \lfloor \log L_i / \log 2 \rfloor$ and $L_{j,i} := L_i^{j/N_i}$, for all $j = 0, \dots, N_i$. Let $\zeta > 0$ and $w > 0$.

We say that a point $x \in M$ is a $(w, (L_i), \zeta)$ -good point for the basis \mathcal{F}_α if having set $y_i = \phi_{X_\alpha}^{L_i}(x)$, for all $i \in \mathbb{N}$ and for all $0 \leq j \leq N_i$, we have

$$w_{\mathcal{F}_\alpha^{(L_{j,i})}}(x, 1) \geq w/L_i^\zeta, \quad w_{\mathcal{F}_\alpha^{(L_{j,i})}}(y_i, 1) \geq w/L_i^\zeta.$$

Remark 5.17. By its definition the set of $(w, (L_i), \zeta)$ -good points for the basis \mathcal{F}_α is saturated by the orbit of the action of the centre $Z(G)$ of the quasi-Abelian nilpotent Lie group G on M . Moreover, if the vector $\rho \in [0, 1]^a$ of scaling exponents vanishes on all vectors of the basis \mathcal{F}_α which belong to the center $Z(\mathfrak{g})$ of the Lie algebra \mathfrak{g} of G , then a point $x \in M$ is $(w, (L_i), \zeta)$ -good for the basis \mathcal{F}_α if and only if its projection $\bar{x} \in M/Z(G)$ is $(w, (L_i), \zeta)$ -good for the projection $\bar{\mathcal{F}}_\alpha$ of the basis \mathcal{F}_α onto the quotient $\mathfrak{g}/Z(\mathfrak{g})$, that is, onto the Lie algebra of $G/Z(G)$. It follows that in this case, the set of $(w, (L_i), \zeta)$ -good points for the basis \mathcal{F}_α is not only invariant under the action of $Z(G)$ on M but its projection onto $M/Z(G)$ is invariant under the action of the center $Z(G/Z(G))$ of $G/Z(G)$ onto $M/Z(G)$.

Lemma 5.18. Let $\zeta > 0$ be fixed and let (L_i) be an increasing sequence of positive real numbers satisfying the condition

$$(102) \quad \Sigma((L_i), \zeta) := \sum_{i \in \mathbb{N}} (\log L_i)^2 L_i^{-\zeta} < +\infty.$$

Let $\rho \in [0, 1]^a$, with $\sum \rho_i = 1$, $\nu \leq 1/|\bar{\rho}|$ and let $\alpha_1 := (\alpha_1^{(1)}, \dots, \alpha_1^{(n)}) \in D_n(\bar{\rho}/|\bar{\rho}|, \nu)$. Then the Lebesgue measure of the complement of the set $\mathcal{G}(w, (L_i), \zeta)$ of $(w, (L_i), \zeta)$ -good points is bounded above as follows: there exists a constant $K := K(a, n, \nu) > 0$ such that

$$\text{meas}(M \setminus \mathcal{G}(w, (L_i), \zeta)) \leq K \Sigma((L_i), \zeta) [1/I(Y)]^a \mathcal{H}(Y, \rho, \alpha) w.$$

Proof. For all $i \in \mathbb{N}$ and for all $j = 0, \dots, N_i$, let

$$\mathcal{S}_{j,i} = \left\{ z \in M : w_{\mathcal{F}_\alpha^{(L_{j,i})}}(z, 1) < L_i^\zeta/w \right\}.$$

By definition we have

$$(103) \quad M \setminus \mathcal{G}(w, (L_i), \zeta) = \bigcup_{i \in \mathbb{N}} \bigcup_{j=0}^{N_i} (\mathcal{S}_{j,i} \cup \phi_{X_\alpha}^{-L_i} \mathcal{S}_{j,i}).$$

By Lemma 5.5 for all $z \in \mathcal{S}_{j,i}$ we have

$$(I/2)^a L_i^\zeta / w < \int_0^1 H_{L_{j,i}}^1 \circ \phi_{X_\alpha}^\tau(z) d\tau = \frac{1}{L_{j,i}} \int_0^{L_{j,i}} H_{L_{j,i}}^1 \circ \phi_{X_\alpha}^\tau(z) d\tau.$$

It follows that

$$\mathcal{S}_{j,i} \subset \mathcal{S}(j,i) := \left\{ z \in M : \sup_{J>0} \frac{1}{J} \int_0^J H_{n_{j,i}}^1 \circ \phi_{X_\alpha}^\tau(z) d\tau > (I/2)^a L_i^\zeta / w \right\}.$$

By the maximal ergodic theorem, the Lebesgue measure $\text{meas}[\mathcal{S}(j,i)]$ of the set $\mathcal{S}(j,i)$ satisfies the inequality

$$\text{meas}[\mathcal{S}(j,i)] \leq (2/I)^a (w/L_i^\zeta) \int_M H_{L_{j,i}}^1(z) dz.$$

For brevity, let $\mathcal{H} := \mathcal{H}(Y, \rho, \nu)$ denote the constant defined in formula (101). By Theorem 5.15, since by hypothesis $\nu \leq 1/|\bar{\rho}|$ and $\alpha_1 \in D_n(\bar{\rho}/|\bar{\rho}|, \nu)$, there exists a constant $K' := K'(a, n, \nu) > 0$ such that the following bound holds:

$$\left| \int_M H_{L_{j,i}}^1(x) dx \right| \leq K' \mathcal{H} (1 + \log L_{j,i}).$$

The definition of the N_i implies

$$(104) \quad N_i \leq \log L_i / \log 2 < N_i + 1.$$

Hence by the definition of $L_{j,i}$, we have $\log L_{j,i} < 2j \log 2$. Thus, for some constant $K'' := K''(a, n, \nu) > 0$, we have

$$\text{meas}(\mathcal{S}_{j,i}) \leq \text{meas}[\mathcal{S}(j,i)] \leq K'' (2/I)^a \mathcal{H} w (1+j) L_i^{-\zeta}.$$

From this, again by (104), it follows that, for some constant $K''' := K'''(a, n, \nu) > 0$,

$$\text{meas} \left(\bigcup_{j=0}^{N_i} (\mathcal{S}_{j,i} \cup \phi_{X_\alpha}^{-L_i} \mathcal{S}_{j,i}) \right) \leq K''' (2/I)^a \mathcal{H} w (\log L_i)^2 L_i^{-\zeta}.$$

By sub-additivity of the Lebesgue measure, we derive the bound

$$\text{meas} \left(\bigcup_{i \in \mathbb{N}} \bigcup_{j=0}^{N_i} (\mathcal{S}_{j,i} \cup \phi_{X_\alpha}^{-L_i} \mathcal{S}_{j,i}) \right) \leq K''' \Sigma((L_i), \zeta) (2/I)^a \mathcal{H} w.$$

By formula (103) the above estimate concludes the proof. \square

6. BOUNDS ON ERGODIC AVERAGES

Let $(\xi, \dots, \tilde{\eta}_i^{(m)}, \dots)$, with $(m, i) \in J$, be the basis (6) defining the lattice Γ , and let A be the analytic subgroup of G of Lie algebra \mathfrak{a} . We denote by $(\xi, \dots, \eta_i^{(m)}, \dots)$ the Jordan basis defined by (8). As usual, for $\alpha := (\alpha_i^{(m)}) \in \mathbb{R}^J$ let $X_\alpha \in \mathfrak{g} \setminus \mathfrak{a}$ be the vector field, introduced in formula (16), given by the formula

$$X_\alpha = \log \left[x^{-1} \exp \left(\sum_{(m,i) \in J} \alpha_i^{(m)} \tilde{\eta}_i^{(m)} \right) \right].$$

The field X_α generates the flow $\{\phi_\alpha^t\} := \{\phi_{X_\alpha}^t\}$ on the nilmanifold $M = \Gamma \backslash G$.

We denote by $\mathcal{F}_{\alpha, \eta}$ the Jordan basis

$$\mathcal{F}_{\alpha, \eta} := (X_\alpha, \dots, \eta_i^{(m)}, \dots).$$

The Hilbert space $L^2(M)$ splits as a direct sum of irreducible sub-representations of G . The irreducible unitary representations occurring in the decomposition of $L_0(M)$ are unitarily equivalent to the representations $\text{Ind}_A^G(\Lambda)$, obtained by inducing from A to G a character $\chi = \exp \iota \Lambda$ whose coordinates $\Lambda(\tilde{\eta}_i^{(m)})$ with respect to the basis $(\tilde{\eta}_i^{(m)})$ are integer multiples of 2π not all zero.

Definition 6.1. We say that a linear form $\Lambda \in \mathcal{O}$ is *integral* if the coefficients $\Lambda(\tilde{\eta}_i^{(m)})$, $(m, i) \in J$, are integer multiples of 2π . We denote by \widehat{M} the set of co-adjoint orbits $\mathcal{O} \subset \mathfrak{a}^*$ of integral linear forms $\Lambda \in \mathfrak{a}^*$.

By Lemma 4.2, irreducible unitary representations induced by the character $\chi = \exp \iota \Lambda$ factor through the filiform group G/G_Λ , where G_Λ is the normal subgroup of G with Lie algebra given by the ideal, already introduced in formula (39),

$$\mathfrak{J}_\Lambda = \bigcap_{i=0}^{k-1} \ker(\Lambda \circ \text{ad}^i(X_\alpha))$$

Remark 6.2. For our goals it is not restrictive to assume that $\Lambda(\tilde{\eta}_{i_m}^{(m)}) \neq 0$ for some $m \in \{1, \dots, n\}$ with $i_m = k$. In fact, suppose that $H_\Lambda \subset L_0^2(M)$ is a sub-representation unitarily equivalent to $\text{Ind}_A^G(\Lambda)$ and that $\Lambda(\tilde{\eta}_{i_m}^{(m)}) = 0$ for all $m \in \{1, \dots, n\}$ with $i_m = k$. By letting $G_{i_m}^{(m)}$ be the subgroup of $Z(G)$ generated by $\tilde{\eta}_{i_m}^{(m)}$, we have that the representation $\text{Ind}_A^G(\Lambda)$ factorises through the quotient group $G' = G/G_{i_m}^{(m)}$ and occurs as a sub-representation in the quasi-Abelian $(k-1)$ -step nilmanifold $M' := G'/\Gamma'$, where Γ' is the lattice $\Gamma G_{i_m}^{(m)}$ of G' .

Let $\widehat{M}_0 \subset \widehat{M} \cap \mathfrak{a}^*$ the subset of all co-adjoint orbits of forms $\Lambda \in \mathfrak{a}^*$ such that $\Lambda(\tilde{\eta}_{i_m}^{(m)}) \neq 0$ for some $m \in \{1, \dots, n\}$ with $i_m = k$. By definition, for every $\mathcal{O} \in \widehat{M}_0$ and every $\Lambda \in \mathcal{O}$, the induced irreducible unitary representation $\pi_\Lambda^{X_\alpha}$ has exactly degree $k-1$. In fact, for any linear functional $\Lambda \in \mathfrak{a}^*$ the degree of the representation $\pi_\Lambda^{X_\alpha}$ only depends on its co-adjoint orbit.

For any $\mathcal{O} \in \widehat{M}_0$, let $H_\mathcal{O}$ denote the primary subspace of $L^2(M)$ which is a direct sum of sub-representations equivalent to $\text{Ind}_A^G(\Lambda)$, for any $\Lambda \in \mathcal{O}$; this space is well-defined since the unitary representations $\text{Ind}_A^G(\Lambda)$ are unitarily equivalent for all $\Lambda \in \mathcal{O}$. For any adapted basis $\mathcal{F} = (X_\alpha, Y)$ and for all $r \in \mathbb{R}$, we set

$$W^r(H_\mathcal{O}, \mathcal{F}) := H_\mathcal{O} \cap W^r(M, \mathcal{F}),$$

which is a Hilbert space once endowed with the transversal Sobolev norm $|\cdot|_{\mathcal{F}, r}$.

For $\mathcal{O} \in \widehat{M}_0$ and $\Lambda \in \mathcal{O}$, let m_0 be any integer such that the vector $\eta_1^{(m_0)}$ has maximal degree k . Then $(X_\alpha, \eta_1^{(m_0)}, \dots, \eta_k^{(m_0)})$ is a k -step filiform basis projecting to a k -step filiform basis of the Lie algebra $\mathfrak{g}/\mathfrak{J}_\Lambda$. By Lemma 4.2 we can complete the system

$$(X_\alpha, Y_1^{(1)}, \dots, Y_k^{(1)}) := (X_\alpha, \eta_1^{(m_0)}, \dots, \eta_k^{(m_0)})$$

to a basis $(X_\alpha, Y_i^{(m)})$ of \mathfrak{g} so that the elements $Y_i^{(m)}$ with $m \neq m_0$ span the ideal \mathfrak{J}_Λ .

Definition 6.3. From now on given a co-adjoint orbit $\mathcal{O} \in \widehat{M}_0$ and a linear functional $\Lambda \in \mathcal{O}$, the symbols $\mathcal{F}_{\alpha, \Lambda}$ and (X_α, Y_Λ) will denote the basis

$$\mathcal{F}_{\alpha, \Lambda} = (X_\alpha, Y_\Lambda) := (X_\alpha, \dots, Y_i^{(m)}, \dots)$$

obtained by completion of the system $(X_\alpha, \eta_1^{(m_0)}, \dots, \eta_k^{(m_0)})$.

Clearly the basis $\mathcal{F}_{\alpha, \Lambda}$ is defined up to an arbitrary choice of the integer m_0 . This ambiguity is irrelevant for what follows; later on we shall make a more precise choice. By construction the basis $\mathcal{F}_{\alpha, \Lambda}$ satisfies the estimates of Lemma 4.2 and it is a normalised, Jordan basis which is also a generalised filiform basis (in the sense of Definition 4.3) for the induced representation $\pi_\Lambda^{X_\alpha}$.

Remark 6.4. The weights introduced in formulas (53) and (69) have a simple expression for the bases $\mathcal{F}_{\alpha, \eta}$ and $\mathcal{F}_{\alpha, \Lambda}$. In fact for any co-adjoint orbit $\mathcal{O} \in \widehat{M}_0$ and for all $\Lambda \in \mathcal{O}$ we have that:

$$|\Lambda(\mathcal{F}_{\alpha, \eta})| = \max_{1 \leq m \leq n} \max_{1 \leq j \leq i_m} |\Lambda(\eta_j^{(m)})|, \quad |\Lambda(\mathcal{F}_{\alpha, \Lambda})| = \max_{1 \leq j \leq k} |\Lambda(\eta_j^{(m_0)})|$$

and

$$\text{for } \mathcal{F} = \mathcal{F}_{\alpha, \eta} \text{ or } \mathcal{F}_{\alpha, \Lambda} \quad \|\Lambda\|_{\mathcal{F}} = |\Lambda(\mathcal{F})| \left(1 + \frac{1}{|\Lambda(\eta_k^{(m_0)})|} \right).$$

From the above formula, it is immediate that for any co-adjoint orbit $\mathcal{O} \in \widehat{M}_0$, for any $\Lambda \in \mathcal{O}$, since $\Lambda(\eta_k^{(m_0)})$ is a non zero integer, we have

$$|\Lambda(\mathcal{F}_{\alpha, \eta})| \leq \|\Lambda\|_{\mathcal{F}_{\alpha, \eta}} \leq 2 |\Lambda(\mathcal{F}_{\alpha, \eta})| \quad \text{and} \quad |\Lambda(\mathcal{F}_{\alpha, \Lambda})| \leq \|\Lambda\|_{\mathcal{F}_{\alpha, \Lambda}} \leq 2 |\Lambda(\mathcal{F}_{\alpha, \Lambda})|.$$

6.1. Coboundary estimates for rescaled bases. In this section we prove Sobolev estimates, with respect to rescaled bases, for the orthogonal projections of the probability measures supported on orbit segments of a quasi-Abelian nilflow on the orthogonal complement of the space of invariant distributions. Our estimates will be derived in every given irreducible unitary representation from the estimates on coboundaries proved in section 4. The rescaled norms will be defined with respect to a generalised filiform basis depending on the irreducible representation.

Let $\rho = (\rho_i^{(m)}) \in (\mathbb{R}^+)^J$ be a fixed vector (to be chosen later) such that

$$\sum_{(m, i) \in J} \rho_i^{(m)} = 1.$$

For all $t \in \mathbb{R}$, let $\mathcal{F}_{\alpha, \Lambda}(t)$ denote the rescaled basis

$$\mathcal{F}_{\alpha, \Lambda}(t) := (X_\alpha(t), Y_\Lambda(t)) = A_t^\rho(X_\alpha, Y_\Lambda) = (e^t X_\alpha, \dots, e^{-t \rho_i^{(m)}} Y_i^{(m)}, \dots).$$

Since the basis (X_α, Y_Λ) is generalised filiform for the induced representation $\pi_\Lambda^{X_\alpha}$, in the optimal choices of the scaling exponents $(\rho_i^{(m)})$ we will always have that

$$\rho_k^{(m_0)} = 0 \quad \text{and} \quad \rho_i^{(m)} = 0, \quad \text{for all } m \neq 1 \text{ and all } i \neq 1.$$

Let us recall that by Definition 4.1, for all $(m, i) \in J$ the degree $d_i^{(m)}$ of the element $Y_i^{(m)} \in \mathfrak{a}$ with respect to the induced representation $\pi_\Lambda^{X_\alpha}$ is the degree of the polynomial $\Lambda(\text{Ad}(e^{x X_\alpha}) Y_i^{(m)})$. Since the basis (X_α, Y_Λ) is generalised filiform

for the induced representation $\pi_\Lambda^{X_\alpha}$, which by construction has maximal degree equal to $k - 1$, it follows that

$$d_i^{(m)} = \begin{cases} k - i, & \text{for } m = m_0 \text{ and for all } i = 1, \dots, i_1 = k; \\ 0, & \text{for } m \neq m_0 \text{ and for all } i = 1, \dots, i_m \leq k. \end{cases}$$

In the present case the exponent $\lambda_{\mathcal{F}}(\rho)$ defined in formula (67) becomes

$$\lambda_{\mathcal{F}}(\rho) = \lambda(\rho) := \min_{1 \leq i < k} \left\{ \frac{\rho_i^{(m_0)}}{k - i} \right\};$$

we also set

$$\delta(\rho) := \min_{1 \leq i < k} \{ \rho_i^{(m_0)} - \rho_{i+1}^{(m_0)} \}.$$

Lemma 6.5. *Let*

$$R^{(m_0)}(\rho) := \sum_{1 \leq i < k} \rho_i^{(m_0)}.$$

We have

$$\delta(\rho) \leq \lambda(\rho) \leq \frac{2R^{(m_0)}(\rho)}{k(k-1)}$$

The above inequalities are both strict unless one has

$$\rho_i^{(m_0)} = \frac{2R^{(m_0)}(\rho)(k-i)}{k(k-1)}, \quad \text{for all } i = 1, \dots, k,$$

in which case they are both equalities.

Lemma 6.6. *There exists a constant $C > 0$ such that, for all $r \in \mathbb{R}^+$ and for any function $f \in W^r(H_\Theta, \mathcal{F}_{\alpha, \Lambda})$ we have*

$$\sum_{(m,i) \in J} |[X_\alpha(t), Y_i^{(m)}(t)]f|_{r, \mathcal{F}_{\alpha, \Lambda}(t)} \leq C e^{t(1-\delta(\rho))} |f|_{r+1, \mathcal{F}_{\alpha, \Lambda}(t)}.$$

Proof. For all $(m, i) \in J^-$, we have $[X_\alpha(t), Y_i^{(m)}(t)] = e^{t(1-\rho_i^{(m)}+\rho_{i+1}^{(m)})} Y_{i+1}^{(m)}(t)$. Since $\Lambda \in \Theta$, by construction we have $Y_i^{(m)} \in \mathfrak{J}_\Lambda$, for $m \neq 1$ and for all $i = 1, \dots, i_m$. Since \mathfrak{J}_Λ coincides with the kernel of the induced representation $\pi_\Lambda^{X_\alpha}$, which is unitarily equivalent to the representation given by the action of G on H_Θ , in the space H_Θ we have

$$[X_\alpha(t), Y_i^{(m)}(t)]f = 0, \quad \text{for all } m \neq 1 \text{ and for all } i = 1, \dots, i_m.$$

It follows that

$$\sum_{(m,i) \in J} |[X_\alpha(t), Y_i^{(m)}(t)]f|_{r, \mathcal{F}_{\alpha, \Lambda}(t)} \leq e^{t(1-\delta(\rho))} \sum_{i=2}^k |Y_i^{(m_0)}(t)f|_{r, \mathcal{F}_{\alpha, \Lambda}(t)},$$

thereby concluding the proof. \square

Proposition 6.7. *Let $r > (a/2 + 1)(k - 1) + 1$ and let $\mathcal{F}_{\alpha, \Lambda}(t)$ and $\rho \in (\mathbb{R}^+)^a$ be defined as above. For $x \in M$ let γ_x be the Birkhoff average operator*

$$(105) \quad \gamma_x(f) = \frac{1}{L} \int_0^L f(\phi_{X_\alpha}^\tau(x)) d\tau$$

and consider the decomposition of the restriction of the linear functional γ_x to $W_0^r(H_\Theta, \mathcal{F}_{\alpha, \Lambda}(t))$ as an orthogonal sum $\gamma_x = D(t) + R(t)$ in $W_0^{-r}(H_\Theta, \mathcal{F}_{\alpha, \Lambda}(t))$ of a X_α -invariant distribution $D(t)$ and an orthogonal complement $R(t)$.

There is a constant $C_r^{(m_0)}$ such that for all $g \in W^r(H_\mathcal{O}, \mathcal{F}_{\alpha, \Lambda}(t))$ and all $t \geq 0$, having set $y = \phi_{X_\alpha}^L(x)$, we have

$$|R(t)(g)| \leq C_r^{(1)} (1 + \|\Lambda\|_{\mathcal{F}_{\alpha, \Lambda}})^{\frac{k(r+1)-2}{k-1}} e^{t(\lambda(\rho) - \delta(\rho))} L^{-1} \\ \times \left(\frac{1}{w_{\mathcal{F}_{\alpha, \Lambda}(t)}(x, 1)^{1/2}} + \frac{1}{w_{\mathcal{F}_{\alpha, \Lambda}(t)}(y, 1)^{1/2}} \right) |g|_{r, \mathcal{F}_{\alpha, \Lambda}(t)}.$$

Proof. Fix $t \geq 0$ and for brevity set $D = D(t)$, $R = R(t)$. Let $g \in W^r(H_\mathcal{O}, \mathcal{F}_{\alpha, \Lambda}(t))$. We write $g = g_D + g_R$, where g_R is in the kernel of the X_α -invariant distributions and g_D is orthogonal to g_R in $W^r(H_\mathcal{O}, \mathcal{F}_{\alpha, \Lambda}(t))$. Then g_R is a coboundary and $R(g_D) = 0$. Let $f := G_{X_\alpha, \Lambda}^{X_\alpha(t)}(g_R)$. From $R(g_D) = 0$ and $D(g_R) = 0$ we obtain

$$|R(g)| = |R(g_D + g_R)| = |R(g_R)| = |\gamma_x(g_R) - D(g_R)| \\ = |\gamma_x(g_R)| \leq \frac{1}{L} (|f(x)| + |f(y)|).$$

By Theorem 3.9 and Lemma 6.6, for any $\tau > a/2 + 1$ there exist positive constants C_τ and C such that for any $z \in M$ we have the estimate

$$|f(z)| \leq \frac{C_\tau}{w_{\mathcal{F}_{\alpha, \Lambda}(t)}(z, 1)^{1/2}} \left(C e^{t(1-\delta(\rho))} |f|_{\tau, \mathcal{F}_{\alpha, \Lambda}(t)} + |g_R|_{\tau-1, \mathcal{F}_{\alpha, \Lambda}(t)} \right).$$

By Theorem 4.14 for $r > \tau(k-1) + 1$ we have, for all $t \geq 0$,

$$|f|_{\tau, \mathcal{F}_{\alpha, \Lambda}(t)} \leq G_{k, r, \tau} (1 + \|\Lambda\|_{\mathcal{F}_{\alpha, \Lambda}})^{\tau k + 2} e^{-(1-\lambda(\rho))t} |g_R|_{r, \mathcal{F}_{\alpha, \Lambda}(t)}.$$

The conclusion follows from the estimates above and the observation that, by orthogonality, we have $|g_R|_{r, \mathcal{F}_{\alpha, \Lambda}(t)} \leq |g|_{r, \mathcal{F}_{\alpha, \Lambda}(t)}$. \square

Corollary 6.8. *For every $r > (a/2 + 1)(k-1) + 1$, there is a constant $C_r^{(2)}$ such that the following holds true for every $\mathcal{O} \in \widehat{M}_0$, every $\Lambda \in \mathcal{O}$ and every $x \in M$. Let γ_x be the Birkhoff average operator (105) and let $\gamma_x = D + R$ be the decomposition of γ_x as an orthogonal sum in $W_0^{-r}(H_\mathcal{O}, \mathcal{F}_{\alpha, \Lambda})$ of an X_α -invariant distribution D and an orthogonal complement R . Then*

$$|R|_{-r, \mathcal{F}_{\alpha, \Lambda}} \leq C_r^{(2)} [1/I(Y_\Lambda)]^{a/2} (1 + \|\Lambda\|_{\mathcal{F}_{\alpha, \Lambda}})^{\frac{k(r+1)-2}{k-1}} L^{-1}.$$

Proof. By our definitions, the width function $x \in M \mapsto w_{\mathcal{F}_{\alpha, \Lambda}}(x, 1)^{-1}$ is uniformly bounded on M . In fact, for all $x \in M$ we have

$$w_{\mathcal{F}_{\alpha, \Lambda}}(x, 1) \geq \left(\frac{I(Y_\Lambda)}{2} \right)^a.$$

The above statement then follows immediately from Proposition 6.7 applied to the orthogonal decomposition $\gamma_x = D(0) + R(0)$ in the Hilbert space $W^r(H_\mathcal{O}, \mathcal{F}(0))$. \square

6.2. Bounds on ergodic averages in a irreducible sub-representation. In this section we derive bounds on ergodic averages of quasi-Abelian nilflows for functions in a single irreducible sub-representation. The proof follows an inductive argument based on the coboundary bounds proved above and on the scaling of invariant distributions proved in section 4.

For brevity, let us set

$$(106) \quad C_r(\Lambda) := (1 + \|\Lambda\|_{\mathcal{F}_{\alpha, \Lambda}})^{\frac{(2k-1)(r+1)-2}{k-1}}.$$

Proposition 6.9. *Let $r > (a/2 + 1)(k - 1) + 1$. Let (L_i) be an increasing sequence of positive real numbers ≥ 1 , let $0 < w \leq I(Y_\Lambda)^a$ and let $\zeta > 0$. There exists a constant $C_r(\rho)$ such that for every $(w, (L_i), \zeta)$ -good point $x \in M$ for the basis $\mathcal{F}_{\alpha, \Lambda}$, for all $i \in \mathbb{N}$ and all $f \in W^r(H_\mathbb{O}, \mathcal{F}_{\alpha, \Lambda})$, we have*

$$(107) \quad \left| \frac{1}{L_i} \int_0^{L_i} f \circ \phi_{X_\alpha}^\tau(x) d\tau \right| \leq C_r(\rho) C_r(\Lambda) w^{-1/2} L_i^{-\delta(\rho) + \lambda(\rho)/2 + \zeta/2} |f|_{r, \mathcal{F}_{\alpha, \Lambda}}.$$

Proof. Let us set $N_i := \lfloor \log L_i / \log 2 \rfloor$ and $t_{j,i} := \log L_{j,i} := \log L_i^{j/N_i}$, for all $j = 0, \dots, N_i$, and observe that

$$(108) \quad N_i \leq \log L_i / \log 2 < N_i + 1, \quad \text{hence} \quad L_i^{1/(N_i+1)} < 2 \leq L_i^{1/N_i} < 4.$$

Let $\mathcal{G}(w, (L_i), \zeta)$ be the set of $(w, (L_i), \zeta)$ -good points. Let us adopt the following notation: for all $t > 0$, $\mathcal{F}_{\alpha, \Lambda}(t) := A_\rho^t \mathcal{F}_{\alpha, \Lambda}$ and, for all $i \in \mathbb{N}$, let $y_i := \phi_{X_\alpha}^{L_i}(x)$. By the definition of a good point (see Definition 5.16), for every good point $x \in \mathcal{G}(w, (L_i), \zeta)$, for all $i \in \mathbb{N}$ and for all $j = 0, \dots, N_i$, we have

$$(109) \quad \frac{1}{w_{\mathcal{F}(t_{j,i})}(x, 1)} \leq L_i^\zeta / w \quad \text{and} \quad \frac{1}{w_{\mathcal{F}(t_{j,i})}(y_i, 1)} \leq L_i^\zeta / w.$$

Let $x \in \mathcal{G}(w, (L_i), \zeta)$ and fix $i \in \mathbb{N}$. For simplicity, we omit the index $i \in \mathbb{N}$ and we set $\gamma(f) = \frac{1}{L_i} \int_0^{L_i} f \circ \phi_{X_\alpha}^\tau(x) d\tau$, $L = L_i$, $N = N_i$, $y = y_i$ and $t_j = t_{j,i}$.

We will also denote by $|\cdot|_{r,j}$ and by $\|\cdot\|_{r,j}$, respectively, the transversal Sobolev norms $|\cdot|_{\mathcal{F}(t_j), r}$ and the transverse Lyapunov-Sobolev norms $\|\cdot\|_{\mathcal{F}(t_j), r}$ relative to the rescaled bases $\mathcal{F}(t_j)$, $j = 0, \dots, N$ (see (74)).

Our goal is to estimate $|\gamma|_{\mathcal{F}_{\alpha, \Lambda}, -r} = |\gamma|_{-r, 0}$. For each $j = 0, \dots, N$, let

$$\gamma = D_j + R_j$$

be the orthogonal decomposition of γ in the Hilbert space $W^{-r}(H_\mathbb{O}, \mathcal{F}(t_j))$ into a X_α -invariant distribution D_j and an orthogonal complement R_j .

By the triangle inequality and Corollary 6.8, we have

$$(110) \quad \begin{aligned} |\gamma|_{-r, 0} &\leq |D_0|_{-r, 0} + |R_0|_{-r, 0} \\ &< |D_0|_{-r, 0} + C_r^{(2)} [1/I(Y_\Lambda)]^{a/2} (1 + \|\Lambda\|_{\mathcal{F}_{\alpha, \Lambda}})^{\frac{k(r+1)-2}{k-1}} L^{-1}. \end{aligned}$$

Thus we turn to estimating $|D_0|_{-r, 0}$. By the definition of the Lyapunov-Sobolev norm (74) and the bounds (75) we have

$$(111) \quad |D_0|_{-r, 0} \leq D_{k,r} (1 + \|\Lambda\|_{\mathcal{F}_{\alpha, \Lambda}})^{r+1} \|D_0\|_{-r, 0}.$$

Let us observe that, since $D_j + R_j = D_{j-1} + R_{j-1}$, we have

$$D_{j-1} = D_j + R'_j,$$

where R'_j denotes the orthogonal projection of R_j , in the space $W^{-r}(H_\mathbb{O}, \mathcal{F}(t_{j-1}))$, on the space of X_α -invariant distributions. It follows that

$$(112) \quad \begin{aligned} \|D_{j-1}\|_{-r, j-1} &\leq \|D_j\|_{-r, j-1} + \|R'_j\|_{-r, j-1} \\ &\leq \|D_j\|_{-r, j-1} + |R'_j|_{-r, j-1} \\ &\leq \|D_j\|_{-r, j-1} + |R_j|_{-r, j-1} \end{aligned}$$

Sub-lemma. *There exists a constant $C := C(r) > 0$ such that, for all $j = 0, \dots, N$,*

$$C^{-1} |\cdot|_{-r, j} \leq |\cdot|_{-r, j-1} \leq C |\cdot|_{-r, j}.$$

Proof of the sub-lemma. Let us observe that $\mathcal{F}(t_j) = A_\rho^{t_j - t_{j-1}} \mathcal{F}(t_{j-1})$ and that, by the inequalities (108), $t_j - t_{j-1} = (\log L)/N \leq 2 \log 2$. Thus in passing from the frame $\mathcal{F}(t_{j-1})$ to the frame $\mathcal{F}(t_j)$ the distortion of the corresponding transversal Sobolev norms (and of their dual norms) is uniformly bounded. \square

By the above sub-lemma, the inequality (112) becomes

$$(113) \quad \|D_{j-1}\|_{-r, j-1} \leq \|D_j\|_{-r, j-1} + C |R_j|_{-r, j}.$$

By Lemma 4.15, with respect to the Lyapunov-Sobolev norms, we have that, for any X_α -invariant distribution D and for all $t \geq s$,

$$\|D\|_{\mathcal{F}(s), -r} \leq e^{-\lambda(\rho)(t-s)/2} \|D\|_{\mathcal{F}_{\alpha, \Lambda}(t), -r},$$

from which, by taking into account that $\mathcal{F}(t_j) = A_\rho^{t_j - t_{j-1}} \mathcal{F}(t_{j-1})$, we obtain

$$\|D_j\|_{-r, j-1} \leq L^{-\lambda(\rho)/2N} \|D_j\|_{-r, j}.$$

Then, setting $\beta = \lambda(\rho)/2N$, from (113) we conclude by finite induction that

$$(114) \quad \begin{aligned} \|D_0\|_{-r, 0} &\leq L^{-\lambda(\rho)/2} \left(\|D_N\|_{-r, N} \right. \\ &\quad \left. + C \sum_{\ell=0}^{N-1} L^{(\ell+1)\beta} |R_{N-\ell}|_{-r, N-\ell} \right) = L^{-\lambda(\rho)/2} (I + II) \end{aligned}$$

Sub-lemma. For any $r > a/2$ there exists a constant $C_r > 0$ such that, for all good points $x \in G(w, (L_i), \zeta)$, we have

$$\|D_N\|_{-r, N} \leq C_r L^{\zeta/2} w^{1/2}.$$

Proof of the sub-lemma. By the definition of Lyapunov-Sobolev norms and by orthogonality we have $\|D_N\|_{-r, N} \leq |D_N|_{-r, N} \leq |\gamma|_{-r, N}$.

The orbit segment $(\phi_{X_\alpha}^\tau(x))_{0 \leq \tau \leq L}$ coincides with the orbit segment $(\phi_{X_\alpha(t_N)}^\tau(x))_{0 \leq \tau \leq 1}$ since $X_\alpha(t_N) = X_\alpha(\log L) = L X_\alpha$. Hence, using the notation of Theorem 3.10, we have $\gamma = B_{X_\alpha}^L(x) = B_{X_\alpha(t_N)}^1(x)$. By that theorem, we have $|\gamma|_{-r, N} = |\gamma|_{\mathcal{F}(t_N), -r} = |B_{X_\alpha(t_N)}^1(x)|_{\mathcal{F}(t_N), -r} \leq C_r w_{\mathcal{F}(t_N)}(x, 1)^{-1/2}$. By the inequality (109) we also have $w_{\mathcal{F}(t_N)}(x, 1)^{-1/2} \leq L^{\zeta/2} w^{1/2}$, thereby proving the statement. \square

Sub-lemma. For every $r > (a/2 + 1)(k-1) + 1$ there is a constant $C_r(\rho)$ such that for all good points $x \in G(w, (L_i), \zeta)$, we have

$$\sum_{\ell=0}^{N-1} L^{(\ell+1)\beta} |R_{N-\ell}|_{-r, N-\ell} \leq C_r(\rho) w^{-1/2} (1 + \|\Lambda\|_{\mathcal{F}_{\alpha, \Lambda}})^{\frac{k(r+1)-2}{k-1}} L^{\zeta/2 + \lambda(\rho) - \delta(\rho)}.$$

Proof of the sub-lemma. The orbit segment $(\phi_{X_\alpha}^\tau(x))_{0 \leq \tau \leq L}$ has length $L^{\ell/N}$ with respect to the generator $X_\alpha(t_{N-\ell}) = X_\alpha((1 - \ell/N) \log L) = L^{1-\ell/N} X_\alpha$. Thus, by

Proposition 6.7, with $e^{t_{N-\ell}(\lambda(\rho)-\delta(\rho))} = L^{(1-\frac{\ell}{N})(\lambda(\rho)-\delta(\rho))}$, we obtain

$$\begin{aligned} |R_{N-\ell}(g)|_{-r, N-\ell} &\leq C_r^{(1)} (1 + \|\Lambda\|_{\mathcal{F}_{\alpha, \Lambda}})^{\frac{k(r+1)-2}{k-1}} L^{(1-\frac{\ell}{N})(\lambda(\rho)-\delta(\rho))-\frac{\ell}{N}} \\ &\quad \times \left(\frac{1}{w_{\mathcal{F}(t_{N-\ell})}(x, 1)^{1/2}} + \frac{1}{w_{\mathcal{F}(t_{N-\ell})}(y, 1)^{1/2}} \right) \\ &\leq 2C_r^{(1)} w^{-1/2} (1 + \|\Lambda\|_{\mathcal{F}_{\alpha, \Lambda}})^{\frac{k(r+1)-2}{k-1}} L^{(1-\frac{\ell}{N})(\lambda(\rho)-\delta(\rho))-\frac{\ell}{N}+\zeta/2}, \end{aligned}$$

where in the last upper bound we used the inequalities (109). Writing, for simplicity, $C := 2C_r^{(1)} w^{-1/2} (1 + \|\Lambda\|_{\mathcal{F}_{\alpha, \Lambda}})^{\frac{k(r+1)-2}{k-1}}$ and recalling that $\beta = \lambda(\rho)/2N$ we obtain

$$\begin{aligned} \sum_{\ell=0}^{N-1} L^{(\ell+1)\beta} |R_{N-\ell}|_{-r, N-\ell} &\leq C L^{\zeta/2+\lambda(\rho)-\delta(\rho)} \sum_{\ell=0}^{N-1} L^{(\ell+1)\beta} L^{-\frac{\ell}{N}(\lambda(\rho)-\delta(\rho))-\frac{\ell}{N}} \\ &\leq C L^{\zeta/2+\lambda(\rho)-\delta(\rho)+\lambda(\rho)/2N} \sum_{\ell=0}^{N-1} L^{-\frac{\ell}{N}(1+\lambda(\rho)/2-\delta(\rho))} \\ &\leq 2^r C L^{\zeta/2+\lambda(\rho)-\delta(\rho)} \sum_{\ell=0}^{\infty} 2^{-\ell(1+\lambda(\rho)/2-\delta(\rho))}, \end{aligned}$$

where we have used the inequalities $2 \leq L^{1/N} < 4$. By Lemma 6.5 we have $1 + \lambda(\rho)/2 - \delta(\rho) > 1/2$, concluding the proof of the sub-lemma. \square

By applying the two previous sub-lemmata to the formula (114), and observing that, by Lemma 6.5, $\delta(\rho) - \lambda(\rho)/2 \leq \lambda(\rho)/2$, we obtain that there exists a constant $C_r^{(1)}(\rho)$ such that

$$\|D_0\|_{-r, 0} \leq C_r^{(1)}(\rho) w^{-1/2} (1 + \|\Lambda\|_{\mathcal{F}_{\alpha, \Lambda}})^{\frac{k(r+1)-2}{k-1}} L^{-\delta(\rho)+\lambda(\rho)/2+\zeta/2}.$$

From (110), (111) and the above we conclude that there exists a constant $C_r^{(2)}(\rho)$ such that, whenever $0 < w \leq I(Y_\Lambda)^a$,

$$|\gamma|_{-r, \mathcal{F}} \leq C_r^{(2)}(\rho) w^{-1/2} (1 + \|\Lambda\|_{\mathcal{F}_{\alpha, \Lambda}})^{\frac{(2k-1)(r+1)-2}{k-1}} L^{-\delta(\rho)+\lambda(\rho)/2+\zeta/2},$$

thereby concluding the proof of the proposition. \square

Notation 6.10. Let

$$\widetilde{M}_0 = \bigcup_{\Lambda \in \widetilde{M}_0} \{\Lambda \in \mathcal{O} \mid \Lambda \text{ integral}\}.$$

Theorem 6.11. Let $\sigma = (\sigma_1, \dots, \sigma_n) \in (0, 1)^n$ be such that $\sigma_1 + \dots + \sigma_n = 1$. For any $\Lambda \in \widetilde{M}_0$, let

$$\sigma_\Lambda := \max\{\sigma_m \mid m = 1, \dots, n, \Lambda(\eta_k^{(m)}) \neq 0\}.$$

Let $\nu \in [1, 1 + (k/2 - 1)\sigma_\Lambda]$. Then for any $r > (a/2 + 1)(k - 1) + 1$, there exists a constant $C_r(\sigma, \nu) > 0$ such that the following holds true. For every $\varepsilon > 0$ there exists a constant $K_\varepsilon(\sigma, \nu) > 0$ such that, for every $\alpha := (\alpha_i^{(m)}) \in \mathbb{R}^a$ such that $\alpha_1 := (\alpha_1^{(1)}, \dots, \alpha_1^{(n)}) \in D_n(\sigma, \nu)$ and for every $w \in (0, I(Y_\Lambda)^a]$ there exists a measurable set $\mathcal{G}_\Lambda(\sigma, \varepsilon, w)$ satisfying the estimate

$$(115) \quad \text{meas}(M \setminus \mathcal{G}_\Lambda(\sigma, \varepsilon, w)) \leq K_\varepsilon(\sigma, \nu) \left(\frac{w}{I(Y_\Lambda)^a} \right) \mathcal{H}(Y_\Lambda, \rho, \alpha),$$

with the property that for every $x \in \mathcal{G}_\Lambda(\sigma, \varepsilon, w)$, for every $f \in W^r(H_\odot, \mathcal{F})$ and every $L \geq 1$ we have

$$\left| \frac{1}{L} \int_0^L f \circ \phi_{X_\alpha}^\tau(x) d\tau \right| \leq \frac{C_r(\sigma, \nu) C_r(\Lambda)}{w^{1/2}} L^{-(1-\varepsilon) \frac{2\sigma_\Lambda}{3(k-1)[(k-2)\sigma_\Lambda+2]}} |f|_{r, \mathcal{F}_{\alpha, \Lambda}}.$$

Furthermore, if $w' < w$ we have $\mathcal{G}_\Lambda(\varepsilon, w, k) \subset \mathcal{G}_\Lambda(\varepsilon, w', k)$.

Proof. It is not restrictive to assume, up to renumbering the coordinates of the vector $(\sigma_1, \dots, \sigma_n)$, that $\sigma_\Lambda = \sigma_1$. Let $\rho = (\dots, \rho_i^{(m)}, \dots)$ be the vector given by the following formulas:

$$\begin{aligned} \rho_1^{(m)} &:= \frac{2\sigma_m}{(k-2)\sigma_1+2}, \quad \text{for all } m = 1, \dots, n; \\ \rho_i^{(1)} &:= \frac{2\sigma_1(k-i)}{(k-1)[(k-2)\sigma_1+2]}, \quad \text{for all } i = 2, \dots, k; \\ \rho_i^{(m)} &:= 0, \quad \text{for all } m \neq 1 \text{ and all } i \neq 1. \end{aligned}$$

We can verify that by the hypothesis and the above definition

$$(116) \quad \lambda(\rho) = \delta(\rho) = \frac{2\sigma_1}{(k-1)[(k-2)\sigma_1+2]}.$$

Let us set $\zeta := 2\delta(\rho)/3 - \lambda(\rho)/3$. It is not restrictive to assume that $\zeta > 0$, otherwise the statement is trivially true, since by the Sobolev embedding theorem any function $f \in W^r(H_\odot, \mathcal{F})$ is (uniformly) bounded.

Let $\varepsilon > 0$ and, for all $i \in \mathbb{N}$, let us set $L_i = i^{(1+\varepsilon)\zeta^{-1}}$. Then there exists a constant $K_\varepsilon(\rho) > 0$ such that

$$\Sigma(w, (L_i), \zeta) = \sum_i (\log L_i)^2 L_i^{-\zeta} \leq K_\varepsilon(\rho).$$

Let $\mathcal{G} = \mathcal{G}_\Lambda(\sigma, \varepsilon, w) := \mathcal{G}(w, (L_i), \zeta)$ be the set of $(w, (L_i), \zeta)$ -good points for the basis $\mathcal{F}_{\alpha, \Lambda}$. The estimate in formula (115) follows from Lemma 5.18 and the last statement of the Theorem from the definition of good points. By Proposition 6.9, for all $x \in \mathcal{G}$ and for every $f \in W^r(H_\odot, \mathcal{F})$ the estimate in formula (107) holds true.

Let $L \in [L_i, L_{i+1}]$. Then

$$\int_0^L f \circ \phi_{X_\alpha}^\tau(x) d\tau = \int_0^{L_i} f \circ \phi_{X_\alpha}^\tau(x) d\tau + \int_{L_i}^L f \circ \phi_{X_\alpha}^\tau(x) d\tau = (I) + (II).$$

For brevity, let $C := C_r(\rho) C_r(\Lambda) / w^{1/2}$. The first term is estimated by formula (107):

$$(I) \leq C L^{1-\delta(\rho)+\lambda(\rho)/2+\zeta/2} |f|_{\sigma, \mathcal{F}_{\alpha, \Lambda}} = C L^{1-2\delta(\rho)/3+\lambda(\rho)/3} |f|_{\sigma, \mathcal{F}_{\alpha, \Lambda}}.$$

For the second term, the statement follows from an elementary estimate. In fact, let us set $\beta := (1+\varepsilon)\zeta^{-1}$ and observe that $\beta^{-1} = \zeta(1+\varepsilon)^{-1} \geq (1-\varepsilon)\zeta$. We have

$$\begin{aligned} (II) &\leq (L - L_i) \|f\|_\infty \leq \beta 2^{\beta-1} L^{1-\beta^{-1}} \|f\|_\infty \\ &\leq C'(\rho) L^{1-(1-\varepsilon)(2\delta(\rho)/3-\lambda(\rho)/3)} \|f\|_{r, \mathcal{F}_{\alpha, \Lambda}}. \end{aligned}$$

By the above estimates on the terms (I) and (II) and by the identities in formula (116) for the exponents $\lambda(\rho)$ and $\delta(\rho)$, the proof is completed. \square

6.3. General bounds on ergodic averages. In this section the bounds on ergodic averages obtained above for functions belonging to a single irreducible sub-representation are generalised to all sufficiently smooth functions. The main idea is to use extra regularity of the datum to obtain estimates that are uniform across all irreducible sub-representations.

For all $\mathcal{O} \in \widehat{M}_0$ and $\Lambda \in \mathcal{O}$, the vector

$$(\Lambda(\tilde{\eta}_{i_1}^{(1)}), \Lambda(\tilde{\eta}_{i_2}^{(2)}), \dots, \Lambda(\tilde{\eta}_{i_n}^{(n)})) = (\Lambda(\eta_{i_1}^{(1)}), \Lambda(\eta_{i_2}^{(2)}), \dots, \Lambda(\eta_{i_n}^{(n)})),$$

which obviously depends only on \mathcal{O} , is integral.

For $\mathcal{O} \in \widehat{M}_0$ we define a canonical $\Lambda_{\mathcal{O}} \in \mathcal{O}$ in the following way. For $\Lambda \in \mathcal{O}$, let

$$|\mathcal{O}| = \max_{m=1, \dots, n} \{|\Lambda(\tilde{\eta}_k^{(m)})| \mid i_m = k\}.$$

By the above remarks $|\mathcal{O}|$ does not depend on the choice of $\Lambda \in \mathcal{O}$ and by the definition of \widehat{M}_0 we have $|\mathcal{O}| \neq 0$. Let $m(\mathcal{O}) \in \{1, \dots, n\}$ be the smallest integer m such that

$$i_m = k, \quad \text{and} \quad |\Lambda(\tilde{\eta}_k^{(m)})| = |\mathcal{O}|.$$

Recall that the basis $\mathcal{F}_{\alpha, \Lambda} = (X_{\alpha}, Y_{\Lambda})$ was defined by the choice of an integer m_0 such that the element $\eta_1^{(m_0)}$ had degree k , i.e. such that $\Lambda(\tilde{\eta}_k^{(m_0)}) \neq 0$.

We shall assume henceforth that $m_0 = m(\mathcal{O})$ making, in this way, a unique choice of the basis $\mathcal{F}_{\alpha, \Lambda}$. After relabelling the elements of the basis η we may also assume that $m_0 = m(\mathcal{O}) = 1$. We shall do so, for simplicity of notation.

We first prove estimates for the constants $I(Y_{\Lambda})$, introduced in Definition 5.3, and the constant $\mathcal{H}(Y_{\Lambda}, \rho, \alpha)$, introduced in formula (101), in terms of the weight $|\Lambda(\mathcal{F}_{\alpha, \eta})|$, introduced in formula (53) (see also Remark 6.4).

From Definition 5.3 and Lemma 4.2 we derive the following estimates:

Lemma 6.12. *For every $\mathcal{O} \in \widehat{M}_0$ and for every $\Lambda \in \mathcal{O}$, we have*

$$I(Y_{\Lambda}) \geq \frac{1}{k} \left(1 + \frac{|\Lambda(\mathcal{F}_{\alpha, \eta})|}{|\mathcal{O}|} \right)^{-k}.$$

Proof. The return time of the flow X_{α} to any orbit of the codimension one Abelian subgroup $A \subset G$ is equal to 1. Hence, by Definition 5.3, we have $I(\eta) = 1/2$ for the basis $\eta := (\eta_i^{(j)})$. From the construction above $\eta_1^{(1)}$ is an element of degree $k-1$ and $\text{ad}^{k-1} \eta_1^{(1)} = \eta_k^{(1)}$. By these observations, the statement follows easily from Definition 5.3 and the estimate of Lemma 4.2 of the coefficients of the matrix of change of bases $C^{\eta, Y_{\Lambda}}$ (since the basis $\mathcal{F}_{\alpha, \eta} = (X_{\alpha}, \eta)$ is Jordan). \square

Lemma 6.13. *For every $\mathcal{O} \in \widehat{M}_0$ and for every $\Lambda \in \mathcal{O}$, we have*

$$I(Y_{\Lambda})^{-a} \mathcal{H}(Y_{\Lambda}, \rho, \alpha) \leq \frac{2}{k^a} C(\alpha_1) (1 + \log C(\alpha_1)) \left(1 + \frac{|\Lambda(\mathcal{F}_{\alpha, \eta})|}{|\mathcal{O}|} \right)^{ak}.$$

Proof. By Definition 5.3 we have $I(\eta) \leq 1/2$ and by the definition (100) of $C(\alpha_1)$ we have $C(\alpha_1) \geq 1$. Then from the definition (101) of the constant $\mathcal{H}(Y_{\Lambda}, \rho, \alpha)$,

using $a > n$, we obtain

$$\begin{aligned}
I(Y_\Lambda)^{-a} \mathcal{H}(Y_\Lambda, \rho, \alpha) &= I(Y_\Lambda)^{-a} \\
&\quad + I(Y_\Lambda)^{-n} C(\alpha_1) \left(1 + \log^+ [I(Y)^{-1}] + \log C(\alpha_1) \right) \\
&\leq C(\alpha_1) \left(1 + \log C(\alpha_1) \right) \\
&\quad \times \left(I(Y_\Lambda)^{-a} + I(Y_\Lambda)^{-n} \log^+ [I(Y)^{-1}] \right) \\
&\leq 2 C(\alpha_1) \left(1 + \log C(\alpha_1) \right) I(Y_\Lambda)^{-a}.
\end{aligned}$$

The lemma now follows from Lemma 6.12. \square

We then construct sets of large measure on which bounds for ergodic integrals hold for functions in each irreducible sub-representation with appropriate constants.

Corollary 6.14. *For given $\mathcal{O} \in \widehat{M}_0$, $\Lambda \in \mathcal{O}$, $w > 0$ and $\varepsilon > 0$, let*

$$w_\Lambda := w \cdot |\Lambda(\mathcal{F}_{\alpha, \eta})|^{-a(k+1)+2-\varepsilon}.$$

For $\sigma, \nu, r, \varepsilon$ and $\alpha \in \mathbb{R}^a$ as in Theorem 6.11 let $\mathcal{G}_\Lambda(\sigma, \varepsilon, w_\Lambda)$ be the set given by that theorem. Then, for every $w > 0$ and $\varepsilon > 0$ the set

$$\mathcal{G}(\sigma, \varepsilon, w) := \bigcap_{\Lambda \in \widetilde{M}_0} \mathcal{G}_\Lambda(\sigma, \varepsilon, w_\Lambda)$$

has measure greater than

$$1 - Cw\varepsilon^{-1}, \quad \text{with } C := k^{-a} K_\varepsilon(\sigma, \nu) C(\alpha_1) (1 + \log C(\alpha_1)).$$

Furthermore, if $w' < w$ we have $\mathcal{G}(\varepsilon, w, k) \subset \mathcal{G}(\varepsilon, w', k)$.

Proof. Recalling that $\Lambda(\eta_k^{(1)}) = |\mathcal{O}|$ and $|\Lambda(\mathcal{F}_{\alpha, \eta})|$ are integral multiples of 2π , by Theorem 6.11, Lemma 6.13 and the definition of w_Λ we have

$$\text{meas}(M \setminus \mathcal{G}_\Lambda(\sigma, \varepsilon, w_\Lambda)) \leq \pi^{-ka} C |\Lambda(\mathcal{F}_{\alpha, \eta})|^{-a-\varepsilon},$$

where $C = 2k^{-a} K_\varepsilon(\sigma, \nu) C(\alpha_1) (1 + \log C(\alpha_1))$. Since the cardinal of integral linear forms $\Lambda \in \widetilde{M}_0$ such that $|\Lambda(\mathcal{F}_{\alpha, \eta})| = 2\pi\ell$ is bounded by $(2\ell)^{a-1}$ we have

$$\begin{aligned}
\sum_{\Lambda \in \widetilde{M}_0} \text{meas}(M \setminus \mathcal{G}_\Lambda(\sigma, \varepsilon, w_\Lambda)) &\leq \\
&2^{-a} C' w \sum_{\ell > 0} \sum_{\Lambda \in \widetilde{M}_0 : |\Lambda(\mathcal{F}_{\alpha, \eta})| = 2\pi\ell} \ell^{-a-\varepsilon} \\
&\leq 2^{-1} C' w \sum_{\ell > 0} \ell^{-1-\varepsilon} < 2^{-1} C' w \varepsilon^{-1}.
\end{aligned}$$

The final statement on the monotonicity of the set $\mathcal{G}(\varepsilon, w, k)$ with respect to $w > 0$ follows from the definition of this set and the analogous statement in Theorem 6.11. \square

To sum up the estimates for ergodic integrals we will bound the constants in our estimates for each irreducible sub-representation in terms of higher norms of the datum. This step can be accomplished by making a particular choice of a linear form in every co-adjoint orbit.

Definition 6.15. For every $\mathcal{O} \in \widehat{M}_0$ we define $\Lambda_{\mathcal{O}}$ as the unique integral linear form $\Lambda \in \mathcal{O}$ such that

$$0 \leq \Lambda(\tilde{\eta}_{k-1}^{(1)}) < |\mathcal{O}|.$$

The existence and uniqueness of $\Lambda_{\mathcal{O}}$ follows immediately from the observations that $\Lambda \circ \text{Ad}(\exp(tX_{\alpha}))(\tilde{\eta}_{k-1}^{(1)}) = \Lambda(\tilde{\eta}_{k-1}^{(1)}) + t|\mathcal{O}|$ and that the form $\Lambda \circ \text{Ad}(\exp(tX_{\alpha}))$ is integral for all integer values of $t \in \mathbb{R}$.

Lemma 6.16. *There exists a constant $C(\Gamma) > 0$ such that on the primary subspace $C^\infty(H_{\mathcal{O}})$ the following estimate holds true:*

$$|\Lambda_{\mathcal{O}}(\mathcal{F}_{\alpha,\eta})| \text{Id} \leq C(\Gamma) (1 + \Delta_{\mathcal{F}_{\alpha,\eta}})^{k/2}.$$

Proof. Let $x_0 = -\Lambda_{\mathcal{O}}(\eta_{k-1}^{(1)})/|\mathcal{O}|$. Then there exists a unique $\Lambda' \in \mathcal{O}$ such that $\Lambda'(\eta_{k-1}^{(1)}) = 0$ given by the formula $\Lambda' = \Lambda_{\mathcal{O}} \circ \text{Ad}(e^{x_0 X_{\alpha}})$. Let us recall that, for any $\Lambda \in \mathfrak{a}^*$, the elements $V \in \mathfrak{a}$ are represented in the representation $\pi_{\Lambda}^{X_{\alpha}}$ as multiplication operators by the polynomials

$$(117) \quad \iota P(\Lambda, V)(x) = \iota \Lambda(\text{Ad}(e^{x X_{\alpha}})V).$$

By the definition of the linear form $\Lambda_{\mathcal{O}} \in \mathcal{O}$, the identity $[X_{\alpha}, \eta_{k-1}^{(1)}] = \eta_k^{(1)}$ immediately implies

$$P(\Lambda', \eta_{k-1}^{(1)})(x) = |\mathcal{O}| x.$$

From (117), we have $P(\Lambda', \text{Ad}(e^{-x X_{\alpha}})V)(x) = \Lambda'(V)$ for all $V \in \mathfrak{a}$, or equivalently,

$$\sum_j \frac{(-x)^j}{j!} P(\Lambda', \text{ad}(X_{\alpha})^j V) = \Lambda'(V), \quad \text{for all } V \in \mathfrak{a};$$

hence, for every element $V \in \mathfrak{a}$ we obtain

$$(118) \quad \begin{aligned} \Lambda'(V) &= \sum_j \frac{(-1)^j}{j!} \left(\frac{P(\Lambda', \eta_{k-1}^{(1)})}{|\mathcal{O}|} \right)^j P(\Lambda', \text{ad}(X_{\alpha})^j V) \\ &= |\mathcal{O}|^{1-k} \sum_j \frac{(-1)^j}{j!} P(\Lambda', \eta_{k-1}^{(1)})^j P(\Lambda', \eta_k^{(1)})^{k-1-j} P(\Lambda', \text{ad}(X_{\alpha})^j V). \end{aligned}$$

Let us recall that, for any $\Lambda \in \mathfrak{a}^*$, the transversal Laplacian for a basis \mathcal{F} in the representation $\pi_{\Lambda}^{X_{\alpha}}$ is the operator of multiplication by the polynomial

$$\Delta_{\Lambda, \mathcal{F}} = \sum_{V \in \mathcal{F}} \pi_{\Lambda}^{X_{\alpha}}(V)^2 = \sum_{V \in \mathcal{F}} P(\Lambda, V)^2,$$

hence for all $(m, j) \in J$, the following bound holds:

$$|P(\Lambda', \eta_j^{(m)})| \leq (1 + \Delta_{\Lambda', \mathcal{F}_{\alpha,\eta}})^{1/2},$$

Since by the identity in formula (118) the constant operators $\Lambda'(\eta_j^{(m)})$ are given by polynomial expressions of degree k in the operators $P(\Lambda', \eta_j^{(m)})$ we obtain the estimate

$$|\Lambda'(\mathcal{F}_{\alpha,\eta})| \text{Id} \leq C_1(\Gamma) (1 + \Delta_{\Lambda', \mathcal{F}_{\alpha,\eta}})^{k/2}.$$

Since the representations $\pi_{\Lambda'}^{X_\alpha}$ and $\pi_{\Lambda_0}^{X_\alpha}$ are unitarily intertwined by the translation operator by x_0 and since constant operators commute with translations, we also have

$$|\Lambda'(\mathcal{F}_{\alpha,\eta})| \text{Id} \leq C_1(\Gamma)(1 + \Delta_{\Lambda_0, \mathcal{F}_{\alpha,\eta}})^{k/2}.$$

Finally, the inequality $0 \leq \Lambda(\tilde{\eta}_{k-1}^{(1)}) < |\mathcal{O}|$ implies that x_0 is bounded by a constant depending only on k . Hence the norms of the linear maps $\text{Ad}(\exp(\pm x_0 X_\alpha))$ are bounded by a constant depending only on k . It follows that $|\Lambda_{\mathcal{O}}(\mathcal{F}_{\alpha,\eta})| \leq C_2(k)|\Lambda'(\mathcal{F}_{\alpha,\eta})|$ and the statement of the lemma follows. \square

Corollary 6.17. *There exists a constant $C'(\Gamma)$ such that for all $\mathcal{O} \in \widehat{M}_0$ and for any sufficiently smooth function $f \in H_{\mathcal{O}}$ we have*

$$C_r(\Lambda_{\mathcal{O}}) w_{\Lambda_{\mathcal{O}}}^{-1/2} |f|_{r, \mathcal{F}_{\alpha, \Lambda}} \leq C'(\Gamma) w^{-1/2} |f|_{r+e, \mathcal{F}_{\alpha, \eta}}$$

where $e = a(k+1)k/2 + k \frac{(2k-1)(\sigma+1)-2}{2(k-1)}$.

Proof. Recall that by the definition (106) we have $C_r(\Lambda_{\mathcal{O}}) = (1 + |\Lambda_{\mathcal{O}}(\mathcal{F}_{\alpha, \Lambda_{\mathcal{O}}})|)^{e_1}$ with $e_1 := \frac{(2k-1)(r+1)-2}{k-1}$. Since $|\Lambda_{\mathcal{O}}(\mathcal{F}_{\alpha, \Lambda_{\mathcal{O}}})| \leq |\Lambda_{\mathcal{O}}(\mathcal{F}_{\alpha, \eta})|$ we have

$$C_r(\Lambda_{\mathcal{O}}) w_{\Lambda_{\mathcal{O}}}^{-1/2} \leq w^{-1/2} (1 + |\Lambda_{\mathcal{O}}(\mathcal{F}_{\alpha, \eta})|)^{e_2}$$

with $e_2 := e_1 + a(k+1)$. By Lemma 6.16 on the space $H_{\mathcal{O}}$ we have

$$(1 + |\Lambda_{\mathcal{O}}(\mathcal{F}_{\alpha, \eta})|)^{e_2} \leq C'(\Gamma)(1 + \Delta_{\mathcal{F}_{\alpha, \eta}})^{e_2 k/2}.$$

\square

We are finally ready to derive global estimates for ergodic integrals.

Proposition 6.18. *Let $r > (a+1)(k-1) + 1$. Let $\sigma = (\sigma_1, \dots, \sigma_n) \in (0, 1)^n$ be a positive vector such that $\sigma_1 + \dots + \sigma_n = 1$. Let us set*

$$\sigma_{\min} := \min\{\sigma_m \mid m = 1, \dots, n, i_m = k\}.$$

Let us assume that $\nu \in [1, 1 + (k/2 - 1)\sigma_{\min}]$ and let $\alpha := (\alpha_i^{(m)}) \in \mathbb{R}^a$ be such that $\alpha_1 := (\alpha_1^{(1)}, \dots, \alpha_1^{(n)}) \in D_n(\sigma, \nu)$. For every $\varepsilon > 0$ and $w > 0$, there exists a measurable set $\mathcal{G}(\sigma, \varepsilon, w)$ satisfying

$$\text{meas}(M \setminus \mathcal{G}(\sigma, \varepsilon, w)) \leq C w \varepsilon^{-1}, \quad \text{with } C := k^{-a} K_\varepsilon(\sigma, \nu) C(\alpha_1) (1 + \log C(\alpha_1)),$$

such that for every $x \in \mathcal{G}(\sigma, w, w)$, for every $f \in W^r(M)$ and every $L \geq 1$ we have

$$(119) \quad \left| \frac{1}{L} \int_0^L f \circ \phi_{X_\alpha}^\tau(x) d\tau \right| \leq w^{-1/2} L^{-(1-\varepsilon) \frac{2\sigma_{\min}}{3(k-1)[(k-2)\sigma_{\min}+2]}} |f|_{r, \mathcal{F}_{\alpha, \eta}}.$$

Furthermore, if $w' < w$ we have $\mathcal{G}(\varepsilon, w, k) \subset \mathcal{G}(\varepsilon, w', k)$.

Proof. We have $\tau := r - ak/2 > (a/2 + 1)(k-1) + 1$. Let $f \in W^\sigma(M, \mathcal{F})$ and let $f = \sum_{\mathcal{O} \in \widehat{M}_0} f_{\mathcal{O}}$ be its orthogonal decomposition onto the primary subspaces $H_{\mathcal{O}}$. Clearly $f_{\mathcal{O}} \in W^\tau(H_{\mathcal{O}}, \mathcal{F})$ and the decomposition is also orthogonal in $W^\tau(H_{\mathcal{O}}, \mathcal{F})$.

Having defined for each $\mathcal{O} \in \widehat{M}_0$ the constant $w_{\Lambda_{\mathcal{O}}}$ as in Corollary 6.14, by the same corollary the set

$$\mathcal{G}(\sigma, \varepsilon, w) := \bigcap_{\mathcal{O} \in \widehat{M}_0} \mathcal{G}_{\Lambda_{\mathcal{O}}}(\sigma, \varepsilon, w_{\mathcal{O}})$$

has measure greater than $1 - Cw\varepsilon^{-1}$, where $C = k^{-a}K_\varepsilon(\sigma, \nu)C(\alpha_1)(1 + \log C(\alpha_1))$, and satisfies the required monotonicity property with respect to $w > 0$.

If $x \in \mathcal{G}(\sigma, \varepsilon, w)$, then by Theorem 6.11 and by Corollary 6.17 the following estimate holds true for every $\mathcal{O} \in \widehat{M}_0 \setminus \widehat{M}_0(x)$ and all $L \geq 1$:

$$\left| \frac{1}{L} \int_0^L f_{\mathcal{O}} \circ \phi_{X_\alpha}^\tau(x) d\tau \right| \leq C_\tau(\sigma, \nu) L^{-(1-\varepsilon) \frac{2\sigma_{min}}{3(k-1)[(k-2)\sigma_{min}+2]}} w^{-1/2} |f_{\mathcal{O}}|_{\tau, \mathcal{F}_{\alpha, \eta}}.$$

For any $\tau > 0$ and any $\varepsilon' > 0$, by Lemma 6.16, we have

$$\begin{aligned} \left| \sum_{\mathcal{O} \in \widehat{M}_0} |f_{\mathcal{O}}|_{\tau, \mathcal{F}_{\alpha, \eta}} \right|^2 &\leq \sum_{\mathcal{O} \in \widehat{M}_0} (1 + |\Lambda_{\mathcal{O}}(\mathcal{F}_{\alpha, \eta})|)^{-a-\varepsilon'} \sum_{\mathcal{O} \in \widehat{M}_0} (1 + |\Lambda_{\mathcal{O}}(\mathcal{F}_{\alpha, \eta})|)^{a+\varepsilon'} |f_{\mathcal{O}}|_{\tau, \mathcal{F}_{\alpha, \eta}}^2 \\ &\leq C(a) |f|_{\tau+(a+\varepsilon')k/2, \mathcal{F}_{\alpha, \eta}}^2 \end{aligned}$$

and the theorem follows by the linearity of ergodic averages after renaming the constants. \square

Theorem 6.19. *Let $r > (a+1)(k-1) + 1$. Let $\sigma = (\sigma_1, \dots, \sigma_n) \in (0, 1)^n$ be a positive vector such that $\sigma_1 + \dots + \sigma_n = 1$. Let*

$$\sigma_{min} := \min\{\sigma_m \mid m = 1, \dots, n, i_m = k\}.$$

Let us assume that $\nu \in [1, 1 + (k/2 - 1)\sigma_{min}]$ and let $\alpha := (\alpha_i^{(m)}) \in \mathbb{R}^a$ be such that $\alpha_1 := (\alpha_1^{(1)}, \dots, \alpha_1^{(n)}) \in D_n(\sigma, \nu)$. For every $\varepsilon > 0$ there exists a full measure measurable set $\mathcal{G}(\sigma, \varepsilon)$ and a measurable function $K_\varepsilon : M \rightarrow \mathbb{R}^+$ with $K \in L^p(M)$ for every $p \in [1, 2[$ such that the following holds. For every $f \in W^r(M)$, for every $x \in \mathcal{G}(\sigma, \varepsilon)$ and every $L \geq 1$ we have

$$(120) \quad \left| \frac{1}{L} \int_0^L f \circ \phi_{X_\alpha}^\tau(x) d\tau \right| \leq K_\varepsilon(x) L^{-(1-\varepsilon) \frac{2\sigma_{min}}{3(k-1)[(k-2)\sigma_{min}+2]}} |f|_{\tau, \mathcal{F}_{\alpha, \eta}}.$$

The set $\mathcal{G}(\sigma, \varepsilon)$ and the function $K_\varepsilon \in L^p(M)$ are invariant under the action of $Z(G)$ on M , moreover the set $\mathcal{G}(\sigma, \varepsilon)/Z(G)$ and the function $K_\varepsilon \in L^p(M/Z(G))$ are well-defined and invariant under the action of $Z(G/Z(G))$ on $M/Z(G)$.

Proof. For $i \in \mathbb{N}^+$ let $w_i := 1/2^i C$ and $\mathcal{G}_i := \mathcal{G}(\sigma, \varepsilon, w_i)$, where $\mathcal{G}(\sigma, \varepsilon, w)$ is the set given by the previous proposition and $C = k^{-a}K_\varepsilon(\sigma, \nu)C(\alpha_1)(1 + \log C(\alpha_1))$. Set $K_\varepsilon(x) := 1/w_i^{1/2}$ if $x \in \mathcal{G}_i \setminus \mathcal{G}_{i-1}$. By Proposition 6.18, the sets \mathcal{G}_i are increasing and satisfy $\text{meas}(M \setminus \mathcal{G}_i) \leq 1/2^i \varepsilon$. Hence the set $\mathcal{G}(\sigma, \varepsilon) := \bigcup_{i \in \mathbb{N}^+} \mathcal{G}_i$ has full measure and the function K is in $L^p(M)$ for every $p \in [1, 2[$. By the same proposition for every $x \in \mathcal{G}(\sigma, \varepsilon)$ and every $f \in W^r(M)$ and every $L \geq 1$ we have

$$(121) \quad \left| \frac{1}{L} \int_0^L f \circ \phi_{X_\alpha}^\tau(x) d\tau \right| \leq K_\varepsilon(x) L^{-(1-\varepsilon) \frac{2\sigma_{min}}{3(k-1)[(k-2)\sigma_{min}+2]}} |f|_{\tau, \mathcal{F}_{\alpha, \eta}}.$$

By Remark 5.17, the stated invariance properties of the set $\mathcal{G}(\sigma, \varepsilon)$ and of the function $K_\varepsilon \in L^p(M)$ under the action of the groups $Z(G)$ and $Z(G/Z(G))$ follow immediately from the above definitions. This concludes the proof. \square

Proof of Theorem 1.1. It follows immediately from the above Theorem 6.19 by choosing $\sigma = (1/n, 1/n, \dots, 1/n)$. \square

A particular case of the above theorem is obtained when the group G is a k -step filiform group Fil_k . Then $n = 1$ and the Lie algebra $\mathfrak{g} = \mathfrak{fil}_k$ is generated by the pair (ξ, η_1) ; the only non-trivial commutation relations are

$$[\xi, \eta_i] = \eta_{i+1}, \quad \text{for } i = 1, \dots, k-1.$$

As usual the formulas (10) define another basis $(\tilde{\eta}_i)$ of the Abelian ideal $\mathfrak{a} = \langle \eta_1, \dots, \eta_k \rangle$ and a lattice Γ_k is defined as in (11).

Let $M(\text{Fil}_k) = \Gamma_k \backslash \text{Fil}_k$ denote the compact manifold obtained in this particular case. For $\alpha = (\alpha_1, \dots, \alpha_k) \in \mathbb{R}^k$ the vector field X_α is now given by

$$X_\alpha := \log \left[\exp(-\xi) \exp \left(\sum_{i=1}^k \alpha_i \tilde{\eta}_i \right) \right].$$

Let us recall that when $n = 1$, by Lemma 5.12, the classical Diophantine condition DC_ν implies the Diophantine condition $D(1, \nu)$. Hence we have:

Theorem 6.20 (Filiform case). *Let $r > k^2$. Let $\nu \in [1, k/2]$ and let $\alpha_1 \in DC_\nu$. For every $\varepsilon > 0$ there exists a full measure measurable set $\mathcal{G}_\varepsilon \subset M(\text{Fil}_k)$ and a measurable function $K_\varepsilon : \mathcal{G}_\varepsilon \rightarrow \mathbb{R}^+$, with $K_\varepsilon \in L^p(M(\text{Fil}_k))$ for every $p \in [1, 2[$, such that for every $x \in \mathcal{G}_\varepsilon$, for every $f \in W^r(M(\text{Fil}_k), \mathcal{F})$ of average zero and for all $L \geq 1$ we have*

$$\left| \frac{1}{L} \int_0^L f \circ \phi_{X_\alpha}^\tau(x) d\tau \right| \leq K_\varepsilon(x) L^{-(1-\varepsilon)\frac{2}{3(k-1)k}} |f|_{r, \mathcal{F}, \alpha, \eta}.$$

The set $\mathcal{G}_\varepsilon \subset M(\text{Fil}_k)$ and the positive function $K_\varepsilon \in L^p(M(\text{Fil}_k))$, defined on \mathcal{G}_ε , are invariant under the action of the centre $Z(\text{Fil}_k)$ of the filiform group Fil_k on $M(\text{Fil}_k)$, moreover the set $\mathcal{G}_\varepsilon/Z(\text{Fil}_k)$ and the function $K_\varepsilon \in L^p(M(\text{Fil}_k)/Z(\text{Fil}_k))$ are well-defined and invariant under the action of the quotient $\text{Fil}_k/Z(\text{Fil}_k)$ on the quotient filiform nilmanifold $M(\text{Fil}_k)/Z(\text{Fil}_k)$.

Proof of Corollary 1.2. We refer to the notation introduced in section 2.3.

Let $\alpha = (\alpha_1, 0, \dots, 0) \in \mathbb{R}^k$. By the above Theorem 6.20 and by Lemma 2.7 we have that if $\alpha_1 \in DC_\nu$, with $\nu \in [1, k/2]$ and $r > k^2$, for any function $f \in H^r(\mathbb{T}_o^k)$ of average zero the following bound holds. There exists a full measure measurable set $\mathcal{G}_\varepsilon \subset \mathbb{T}^{k-2}$ and a measurable function $K_\varepsilon : \mathbb{T}^{k-2} \rightarrow \mathbb{R}^+$, with $K_\varepsilon \in L^p(\mathbb{T}^{k-2})$ for every $p \in [1, 2[$, such that for all $(s_1, \dots, s_{k-2}) \in \mathcal{G}_\varepsilon$ and for all $N \geq 1$, we have

$$\left| \sum_{\ell=0}^{N-1} f(P_k(\alpha, \mathbf{s}, \ell)) \right| \leq K_\varepsilon(s_1, \dots, s_{k-2}) N^{1 - \frac{2}{3k(k-1)} + \varepsilon} |f|_{r, \mathcal{F}}.$$

By Lemma 2.5 we see that the coefficients $a_0, a_1, a_2, \dots, a_{k-1}$ of the polynomial

$$P_k(\alpha, \mathbf{s}, N) = \sum_{j=0}^k a_j N^j$$

are linear functions of the coordinates $(s_1, \dots, s_k) \in \mathbb{T}^k$. In particular, as the $(k-2)$ -tuple (s_1, \dots, s_{k-2}) ranges in a set of full measure $\mathcal{G}_\varepsilon \subset \mathbb{T}^{k-2}$ the coefficients a_2, \dots, a_{k-1} of the polynomial $P_k(\alpha, \mathbf{s}, N)$ also range in a subset of full measure of \mathbb{T}^{k-2} , while for every fixed $(k-2)$ -tuple (s_1, \dots, s_{k-2}) as the pair (s_{k-1}, s_k) ranges over all \mathbb{T}^2 , the pair of coefficients (a_0, a_1) also ranges over all \mathbb{T}^2 . \square

REFERENCES

- [FF06] Livio Flaminio and Giovanni Forni, *Equidistribution of nilflows and applications to theta sums*, Ergodic Theory Dynam. Systems **26** (2006), no. 2, 409–433. MR 2218767 (2007c:37003)
- [FF07] ———, *On the cohomological equation for nilflows*, J. Mod. Dyn. **1** (2007), no. 1, 37–60. MR 2261071 (2008h:37003)
- [Fur81] H. Furstenberg, *Recurrence in ergodic theory and combinatorial number theory*, Princeton University Press, Princeton, N.J., 1981, M. B. Porter Lectures. MR 603625 (82j:28010)
- [GT12] B. Green and T. Tao, *The quantitative behaviour of polynomial orbits on nilmanifolds*, Annals of Math. **175** (2012), 465–540.
- [Woo12] T. D. Wooley, *Vinogradov’s mean value theorem via efficient congruencing*, Annals of Math. **175** (2012), 1575–1627.

MATHÉMATIQUES, UNIVERSITÉ LILLE 1, F59655 VILLENEUVE D’ASQ CEDEX, FRANCE

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MARYLAND, COLLEGE PARK, MD USA

E-mail address: livio.flaminio@math.univ-lille1.fr

E-mail address: gforni@math.umd.edu